(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 115151641 A
(43)申请公布日 2022.10.04

(21)申请号 202080096562.4

(22)申请日 2020.12.10

(30)优先权数据
62/947,768 2019.12.13 US
62/947,605 2019.12.13 US

(85)PCT国际申请进入国家阶段日
2022.08.12

(86)PCT国际申请的申请数据
PCT/US2020/064159 2020.12.10

(87)PCT国际申请的公布数据
WO2021/119226 EN 2021.06.17

(71)申请人 阿尔尼拉姆医药品有限公司
地址 美国马萨诸塞州
申请人 里珍纳龙药品有限公司

(72)发明人 E·费施勒维克 S·米尔斯坦
K·布朗 T·齐默尔曼
J·D·麦金因克 D·弗伦德威
E·乔 A·沙玛-坎宁
A·加格里亚蒂 G·德罗古特
B·杜博斯 B·赞布洛维茨

(74)专利代理机构 北京市金杜律师事务所 11256
专利代理师 陈文平 黄海波

(51)Int.Cl.
C12N 15/113(2006.01)
A61K 31/713(2006.01)
A61P 25/28(2006.01)

权利要求书8页 说明书221页 附图47页

(54)发明名称

人类染色体9开放阅读框72(C9ORF72)iRNA剂组合物及其使用方法

(57)摘要

本公开涉及靶向人类染色体9开放阅读框72(C9orf72)基因的双链核酸(dsRNAi)剂及组合物,以及使用此类dsRNAi剂及组合物来抑制C9orf72基因的表达的方法及治疗患有C9orf72相关疾病或疾患的受试者的方法,该疾病或疾患是,例如,C9orf72肌萎缩侧索硬化/额颞叶痴呆或由于C9orf72扩增引起的亨廷顿样综合征。

CN 115151641 A

1.一种用于抑制C9orf72的表达的双链核糖核酸(dsRNA)剂,其中,该dsRNA剂包含形成双链区域的正义链及反义链,

其中,该正义链包含与表10A、10B、12至15、19、20或21任一者中的反义核苷酸序列的任一个相差不超过3个核苷酸的至少15个连续核苷酸。

2.一种用于抑制C9orf72的表达的双链核糖核酸(dsRNA)剂,其中,该dsRNA剂包含形成双链区域的正义链及反义链,

其中,该正义链包含与SEQ ID NO:133的核苷酸200-290相差不超过三个核苷酸的至少15个连续核苷酸,且该反义链包含来自SEQ ID NO:134的相应核苷酸序列的至少15个连续核苷酸,以及

其中,该正义链、该反义链、或该正义链及该反义链二者缀合至一个或多个亲脂性部分。

3.一种用于抑制C9orf72的表达的双链核糖核酸(dsRNA)剂,其中,该dsRNA剂包含形成双链区域的正义链及反义链,

其中,该反义链包含与SEQ ID NO:125的核苷酸序列相差不超过3个核苷酸的至少15个连续核苷酸,以及

其中,该正义链、该反义链、或该正义链及该反义链二者缀合至一个或多个亲脂性部分。

4.如权利要求1或2所述的dsRNA剂,其中,该正义链及反义链的核苷酸序列包含表14、15、19或20任一者中的正义链及反义链核苷酸序列的任一个。

5.如权利要求1、2及4中任一项所述的dsRNA剂,其中,该反义链包含与选自下列所组成的组的双链体的反义链核苷酸序列的任一个相差不超过三个核苷酸的至少15个连续核苷酸:AD-348904.1、AD-348136.1、AD-347612.1、AD-348639.1、AD-347606.1、AD-348905.1、AD-347613.1、AD-348842.1、AD-347610.1、AD-350329.1、AD-348140.1、AD-348930.1、AD-347863.1、AD-348500.1、AD-348597.1、AD-348578.1、AD-347923.1、AD-348588.1、及AD-347773.1。

6.一种用于选择性地抑制C9orf72靶标RNA的双链核糖核酸(dsRNA)剂,该靶标RNA包含六核苷酸重复,该重复包含SEQ ID NO:1的多个连续拷贝,

其中,该dsRNA剂包含形成双链区域的正义链及反义链,

其中,该反义链包含与表14、15及21任一者中的反义核苷酸序列的任一个相差不超过3个核苷酸的至少15个连续核苷酸。

7.如权利要求1、2、4、5及6中任一项所述的dsRNA剂,其中,该反义链包含与选自下列所组成的组的双链体的反义链核苷酸序列的任一个相差不超过三个核苷酸的至少15个连续核苷酸:AD-463863.1、AD-463862.1、AD-463869.1、AD-463873.1、AD-463872.1及AD-463860.1。

8.如权利要求7所述的dsRNA,其中,该反义链包含来自选自下列所组成的组的双链体的反义链核苷酸序列的任一个的至少15个连续核苷酸:AD-463863.1、AD-463862.1、AD-463869.1、AD-463873.1、AD-463872.1及AD-463860.1。

9.如权利要求7所述的dsRNA剂,其中,该反义链包含与选自下列所组成的组的双链体的反义链核苷酸序列的任一个相差不超过三个核苷酸的核苷酸序列:AD-463863.1、AD-

463862.1、AD-463869.1、AD-463873.1、AD-463872.1及AD-463860.1。

10.如权利要求7所述的dsRNA剂，其中，该反义链包含选自下列所组成的组的双链体的反义链核苷酸序列的任一个的核苷酸序列：AD-463863.1、AD-463862.1、AD-463869.1、AD-463873.1、AD-463872.1及AD-463860.1。

11.如权利要求7所述的dsRNA剂，其中，该反义链由选自下列所组成的组的双链体的反义链核苷酸序列的任一个的核苷酸序列组成：AD-463863.1、AD-463862.1、AD-463869.1、AD-463873.1、AD-463872.1及AD-463860.1。

12.如权利要求7至11中任一项所述的dsRNA剂，其中，该正义链包含与选自下列所组成的组的双链体的正义链核苷酸序列的任一个相差不超过三个核苷酸的至少15个连续核苷酸：AD-463863.1、AD-463862.1、AD-463869.1、AD-463873.1、AD-463872.1及AD-463860.1。

13.如权利要求7至12中任一项所述的dsRNA剂，其中，该正义链包含来自选自下列所组成的组的双链体的正义链核苷酸序列的任一个的至少15个连续核苷酸：AD-463863.1、AD-463862.1、AD-463869.1、AD-463873.1、AD-463872.1及AD-463860.1。

14.如权利要求7至13中任一项所述的dsRNA剂，其中，该正义链包含与选自下列所组成的组的双链体的正义链核苷酸序列的任一个相差不超过三个核苷酸的核苷酸序列：AD-463863.1、AD-463862.1、AD-463869.1、AD-463873.1、AD-463872.1及AD-463860.1。

15.如权利要求7至14中任一项所述的dsRNA剂，其中，该正义链包含选自下列所组成的组的双链体的正义链核苷酸序列的任一个的核苷酸序列：AD-463863.1、AD-463862.1、AD-463869.1、AD-463873.1、AD-463872.1及AD-463860.1。

16.如权利要求7至15中任一项所述的dsRNA剂，其中，该正义链由选自下列所组成的组的双链体的正义链核苷酸序列的任一个的核苷酸序列组成：AD-463863.1、AD-463862.1、AD-463869.1、AD-463873.1、AD-463872.1及AD-463860.1。

17.如权利要求1或3所述的dsRNA剂，其中，该正义链及反义链的核苷酸序列包含表12或13任一者中的正义链及反义链核苷酸序列的任一个。

18.如权利要求1、3及17中任一项所述的dsRNA剂，其中，该正义链包含与SEQ ID NO:121的核苷酸230至270、233至262、800至840、800至830、802至828、1240至1290、1240至1280、1247至1288、1590至1645、1590至1620、及1594至1642中任一核苷酸序列相差不超过三个核苷酸的至少15个连续核苷酸，并且该反义链包含来自SEQ ID NO:125的相应核苷酸序列的至少15个连续核苷酸。

19.如权利要求1、3、17及18中任一项所述的dsRNA剂，其中，该正义链包含与SEQ ID NO:121的核苷酸1594至1616、802至824、239至261、1308至1330、233至255、1595至1617、240至262、1532至1554、237至259、3268至3290、806至828、1620至1642、526至548、1169至1191、1266至1288、1247至1269、586至608、1257至1279、及400至422中任一核苷酸序列相差不超过三个核苷酸的至少15个连续核苷酸，并且该反义链包含来自SEQ ID NO:125的相应核苷酸序列的至少15个连续核苷酸。

20.如权利要求1及4至19中任一项所述的dsRNA剂，其中，该正义链、该反义链、或该正义链及该反义链二者缀合至一个或多个亲脂性部分。

21.如权利要求20所述的dsRNA剂，其中，该亲脂性部分缀合至该dsRNA剂的该双链区域中的一个或多个内部位置。

22.如权利要求20或21的所述的dsRNA剂,其中,该亲脂性部分是经由接头或载剂缀合。

23.如权利要求20至22中任一项所述的dsRNA剂,其中,通过logKow测量,该亲脂性部分的亲脂性是超过0。

24.如权利要求1至23中任一项所述的dsRNA剂,其中,通过该双链RNAi剂的血浆蛋白结合测定中的未结合部分测量,该双链RNAi剂的疏水性是超过0.2。

25.如权利要求24所述的dsRNA剂,其中,该血浆蛋白结合测定是使用人血清白蛋白蛋白质的电泳迁移率变动分析。

26.如权利要求1至25中任一项所述的dsRNA剂,其中,该dsRNA剂包含至少一个经修饰的核苷酸。

27.如权利要求26所述的dsRNA剂,其中,该正义链的不超过五个核苷酸及该反义链的不超过五个核苷酸是未经修饰的核苷酸。

28.如权利要求26所述的dsRNA剂,其中,该正义链的全部核苷酸及该反义链的全部核苷酸是经修饰的核苷酸。

29.如权利要求26至28中任一项所述的dsRNA剂,其中,该经修饰的核苷酸的至少一个是选自由下列所组成的组:脱氧核苷酸、3'-端脱氧胸腺嘧啶(dT)核苷酸、2'-O-甲基修饰的核苷酸、2'-氟修饰的核苷酸、2'-脱氧修饰的核苷酸、锁核苷酸、未锁核苷酸、构象限制性核苷酸、约束的乙基核苷酸、无碱基的核苷酸、2'-氨基修饰的核苷酸、2'-O-烯丙基修饰的核苷酸、2'-C-烷基修饰的核苷酸、2'-羟基修饰的核苷酸、2'-甲氧基乙基修饰的核苷酸、2'-O-烷基修饰的核苷酸、吗啉代核苷酸、氨基磷酸酯、包含非天然碱基的核苷酸、四氢呋喃修饰的核苷酸、1,5-失水己糖醇修饰的核苷酸、环己基修饰的核苷酸、包含5'-硫代磷酸酯基团的核苷酸、包含5'-甲基膦酸酯基团的核苷酸、包含5'-磷酸酯或5'-磷酸酯模拟物的核苷酸、包含磷酸乙烯酯的核苷酸、包含腺苷-二醇核酸(GNA)的核苷酸、包含胸苷二醇核酸(GNA)S异构体的核苷酸、包含2-羟甲基-四氢呋喃-5-磷酸酯的核苷酸、包含2'-脱氧胸苷-3'-磷酸酯的核苷酸、包含2'-脱氧鸟苷-3'-磷酸酯的核苷酸、连接至胆固醇基衍生物的末端核苷酸、及十二酸双癸基酰胺基团;及其组合。

30.如权利要求29所述的dsRNA剂,其中,该经修饰的核苷酸是选自由下列所组成的组:2'-脱氧-2'-氟修饰的核苷酸、2'-脱氧修饰的核苷酸、3'-末端脱氧-胸腺嘧啶核苷酸(dT)、锁核苷酸、无碱基的核苷酸、2'-氨基修饰的核苷酸、2'-烷基修饰的核苷酸、吗啉代核苷酸、氨基磷酸酯、以及包含非天然碱基的核苷酸。

31.如权利要求29所述的dsRNA,其中,该经修饰的核苷酸包含3'-末端脱氧-胸腺嘧啶核苷酸(dT)的短序列。

32.如权利要求29所述的dsRNA剂,其中,该核苷酸的修饰是2'-O-甲基修饰、GNA修饰及2'-氟修饰。

33.如权利要求1至32中任一项所述的dsRNA剂,其进一步包含至少一个硫代磷酸酯核苷酸间键。

34.如权利要求33所述的dsRNA剂,其中,该dsRNA剂包含6至8个硫代磷酸酯核苷酸间键。

35.如权利要求1至34中任一项所述的dsRNA剂,其中,每一链是不超过30个核苷酸的长度。

36.如权利要求1至35中任一项所述的dsRNA剂,其中,至少一链包含至少1个核苷酸的3'突出端。

37.如权利要求1至36中任一项所述的dsRNA剂,其中,至少一链包含至少2个核苷酸的3'突出端。

38.如权利要求1至37中任一项所述的dsRNA剂,其中,该双链区域是15至30个核苷酸对的长度。

39.如权利要求38所述的dsRNA剂,其中,该双链区域是17至23个核苷酸对的长度。

40.如权利要求38所述的dsRNA剂,其中,该双链区域是17至25个核苷酸对的长度。

41.如权利要求38所述的dsRNA剂,其中,该双链区域是23至27个核苷酸对的长度。

42.如权利要求38所述的dsRNA剂,其中,该双链区域是19至21个核苷酸对的长度。

43.如权利要求38所述的dsRNA剂,其中,该双链区域是21至23个核苷酸对的长度。

44.如权利要求1至43中任一项所述的dsRNA剂,其中,每一链是19至30个核苷酸的长度。

45.如权利要求1至43中任一项所述的dsRNA剂,其中,每一链是19至23个核苷酸的长度。

46.如权利要求1至43中任一项所述的dsRNA剂,其中,每一链是21至23个核苷酸的长度。

47.如权利要求21至46中任一项所述的dsRNA剂,其中,一个或多个亲脂性部分缀合到至少一链的一个或多个内部位置。

48.如权利要求47的所述的dsRNA剂,其中,该一个或多个亲脂性部分经由接头或载剂缀合到至少一链的一个或多个内部位置。

49.如权利要求48所述的dsRNA剂,其中,该内部位置包括来自该至少一链的每一端的末端两个位置以外的所有位置。

50.如权利要求48所述的dsRNA剂,其中,该内部位置包括来自该至少一链的每一端的末端三个位置以外的所有位置。

51.如权利要求48至50中任一项所述的dsRNA剂,其中,该内部位置排除该正义链的裂解位点区域。

52.如权利要求51所述的dsRNA剂,其中,该内部位置包括从该正义链的5'-端起计数的位置9至12以外的所有位置。

53.如权利要求51所述的dsRNA剂,其中,该内部位置包括从该正义链的3'-端起计数的位置11至13以外的所有位置。

54.如权利要求48至50中任一项所述的dsRNA剂,其中,该内部位置排除该反义链的裂解位点区域。

55.如权利要求54所述的dsRNA剂,其中,该内部位置包括从该反义链的5'-端起计数的位置12至14以外的所有位置。

56.如权利要求48至50中任一项所述的dsRNA剂,其中,该内部位置包括该正义链从3'-端起计数的位置11至13以及该反义链从5'-端起计数的位置12至14以外的所有位置。

57.如权利要求21至56中任一项所述的dsRNA剂,其中,该一个或多个亲脂性部分缀合至一个或多个选自由下列所组成的组的内部位置:该正义链的位置4至8及13至18,以及该

反义链的位置6至10及15至18，自每一链的5’端计数。

58.如权利要求57所述的dsRNA剂，其中，该一个或多个亲脂性部分缀合至一个或多个选自由下列所组成的组的内部位置：该正义链的位置5、6、7、15及17，以及该反义链的位置15及17，自每一链的5’端计数。

59.如权利要求21所述的dsRNA剂，其中，该双链区域中的这些内部位置排除该正义链的裂解位点区域。

60.如权利要求20至59中任一项所述的dsRNA剂，其中，该正义链是21个核苷酸的长度，该反义链是23个核苷酸的长度，以及，该亲脂性部分缀合至该正义链的位置21、位置20、位置15、位置1、位置7、位置6或位置2或该反义链的位置16。

61.如权利要求60所述的dsRNA剂，其中，该亲脂性部分缀合至该正义链的位置21、位置20、位置15、位置1或位置7。

62.如权利要求60所述的dsRNA剂，其中，该亲脂性部分缀合至该正义链的位置21、位置20或位置15。

63.如权利要求60所述的dsRNA剂，其中，该亲脂性部分缀合至该正义链的位置20或位置15。

64.如权利要求60所述的dsRNA剂，其中，该亲脂性部分缀合至该反义链的位置16。

65.如权利要求20至64中任一项所述的dsRNA剂，其中，该亲脂性部分是脂族、脂环族或多脂环族化合物。

66.如权利要求65的所述的dsRNA剂，其中，该亲脂性部分是选自由下列所组成的组：脂质、胆固醇、视黄酸、胆酸、金刚烷乙酸、1-芘丁酸、双氢睾酮、1,3-双-O(十六烷基)甘油、香叶基氧己醇、十六烷基甘油、冰片、薄荷醇、1,3-丙二醇、十七烷基基团、棕榈酸、肉豆蔻酸、O3-(油酰基)石胆酸、O3-(油酰基)胆烯酸、二甲氧基三苯甲基或吩噁嗪。

67.如权利要求65所述的dsRNA剂，其中，该亲脂性部分含有饱和或不饱和的C4-C30烃链，以及选自由羟基、胺、羧酸、磺酸酯、磷酸酯、硫醇、叠氮化物及炔所组成的组的任选的官能基团。

68.如权利要求67所述的dsRNA剂，其中，该亲脂性部分含有饱和或不饱和的C6-C18烃链。

69.如权利要求67所述的dsRNA剂，其中，该亲脂性部分含有饱和或不饱和的C16烃链。

70.如权利要求69所述的dsRNA剂，其中，该饱和或不饱和的C16烃链缀合至自该链的5’-端计数的位置6。

71.如权利要求20至70中任一项所述的dsRNA剂，其中，该亲脂性部分经由载剂缀合，该载剂替换该内部位置或该双链区域中的一个或多个核苷酸。

72.如权利要求71所述的dsRNA剂，其中，该载剂是选自由下列所组成的组的环状基团：吡咯烷基、吡唑啉基、吡唑烷基、咪唑啉基、咪唑烷基、哌啶基、哌嗪基、[1,3]二氧杂环戊基、噁唑烷基、异噁唑烷基、吗啉基、噻唑烷基、异噻唑烷基、喹喔啉基、哒嗪酮基、四氢呋喃基及十氢萘基；或是基于丝氨醇主链或二乙醇胺主链的非环状部分。

73.如权利要求20至70中任一项所述的dsRNA剂，其中，该亲脂性部分经由接头缀合至该双链RNAi剂，该接头含有醚、硫醚、脲、碳酸酯、胺、酰胺、马来酰亚胺-硫醚、二硫化物、磷酸二酯、磺酰胺键、点击反应的产物或氨基甲酸酯。

74.如权利要求20至73中任一项所述的双链iRNA剂,其中,该亲脂性部分缀合至核碱基、糖部分或核苷间键。

75.如权利要求20至74中任一项所述的dsRNA剂,其中,该亲脂性部分或靶向配体经由生物可裂解的接头缀合,该接头是选自由下列所组成的组:DNA,RNA,二硫化物,酰胺,半乳糖胺、葡糖胺、葡萄糖、半乳糖、甘露糖的官能化的单糖或寡糖,及其组合。

76.如权利要求20至75中任一项所述的dsRNA剂,其中,该正义链的3'端是经由端帽保护,该端帽是具有胺的环状基团,该环状基团是选自由下列所组成的组:吡咯烷基、吡唑啉基、吡唑烷基、咪唑啉基、咪唑烷基、哌啶基、哌嗪基、[1,3]二氧杂环戊基、噁唑烷基、异噁唑烷基、吗啉基、噻唑烷基、异噻唑烷基、喹喔啉基、哒嗪酮基、四氢呋喃基及十氢萘基。

77.如权利要求1至19中任一项所述的dsRNA剂,其进一步包含靶向肝脏组织的靶向配体。

78.如权利要求77所述的dsRNA剂,其中,该靶向配体是GalNAc缀合物。

79.如权利要求1至78中任一项所述的dsRNA剂,其进一步包含

末端手性修饰,其出现于该反义链3'端的第一个核苷酸间键处,具有Sp构型的键磷原子,

末端手性修饰,其出现于该反义链5'端的第一个核苷酸间键处,具有Rp构型的键磷原子,以及

末端手性修饰,其出现于该正义链5'端的第一个核苷酸间键处,具有Rp构型或Sp构型的键磷原子。

80.如权利要求1至78中任一项所述的dsRNA剂,其进一步包含

末端手性修饰,其出现于该反义链3'端的第一个及第二个核苷酸间键处,具有Sp构型的键磷原子,

末端手性修饰,其出现于该反义链5'端的第一个核苷酸间键处,具有Rp构型的键磷原子,以及

末端手性修饰,其出现于该正义链5'端的第一个核苷酸间键处,具有Rp或Sp构型的键磷原子。

81.如权利要求1至78中任一项所述的dsRNA剂,其进一步包含

末端手性修饰,其出现于该反义链3'端的第一个、第二个及第三个核苷酸间键处,具有Sp构型的键磷原子,

末端手性修饰,其出现于该反义链5'端的第一个核苷酸间键处,具有Rp构型的键磷原子,以及

末端手性修饰,其出现于该正义链5'端的第一个核苷酸间键处,具有Rp或Sp构型的键磷原子。

82.如权利要求1至78中任一项所述的dsRNA剂,其进一步包含

末端手性修饰,其出现于该反义链3'端的第一个及第二个核苷酸间键处,具有Sp构型的键磷原子,

末端手性修饰,其出现于该反义链3'端的第三个核苷酸间键处,具有Rp构型的键磷原子,

末端手性修饰,其出现于该反义链5'端的第一个核苷酸间键处,具有Rp构型的键磷原

子,以及

末端手性修饰,其出现于该正义链5’端的第一个核苷酸间键处,具有Rp或Sp构型的键磷原子。

83.如权利要求1至78中任一项所述的dsRNA剂,其进一步包含

末端手性修饰,其出现于该反义链3’端的第一个及第二个核苷酸间键处,具有Sp构型的键磷原子,

末端手性修饰,其出现于该反义链5’端的第一个及第二个核苷酸间键处,具有Rp构型的键磷原子,以及

末端手性修饰,其出现于该正义链5’端的第一个核苷酸间键处,具有Rp或Sp构型的键磷原子。

84.如权利要求1至83中任一项所述的dsRNA剂,其进一步包含位于该反义链5’端的磷酸酯或磷酸酯模拟物。

85.如权利要求84所述的dsRNA剂,其中,该磷酸酯模拟物是5’-乙烯基膦酸酯(VP)。

86.如权利要求1至83中任一项所述的dsRNA剂,其中,位于该双链体的反义链5’-末端第1位置处的碱基对是AU碱基对。

87.如权利要求1至83中任一项所述的dsRNA剂,其中,该正义链具有总计21个核苷酸,且该反义链具有总计23个核苷酸。

88.一种细胞,其含有如权利要求1至87中任一项所述的dsRNA剂。

89.一种用于抑制C9orf72的表达的药物组合物,其包含如权利要求1至87中任一项所述的dsRNA剂。

90.一种药物组合物,其包含如权利要求1至87中任一项所述的dsRNA剂,以及脂质制剂。

91.如权利要求89或90所述的药物组合物,其中,该dsRNA剂处于未经缓冲溶液中。

92.如权利要求88所述的药物组合物,其中,该未经缓冲溶液是盐水或水。

93.如权利要求89或90的所述的药物组合物,其中,该dsRNA剂处于缓冲溶液中。

94.如权利要求93所述的药物组合物,其中,该缓冲溶液包含醋酸盐、柠檬酸盐、醇溶谷蛋白、碳酸盐、或磷酸盐、或其任何组合。

95.如权利要求93所述的药物组合物,其中,该缓冲溶液是磷酸盐缓冲盐水(PBS)。

96.一种抑制C9orf72在细胞中的表达的方法,该方法包括使该细胞与如权利要求1至87中任一项所述的dsRNA剂接触,或与如权利要求89至95中任一项所述的药物组合物接触,从而抑制C9orf72在所述细胞中的表达。

97.如权利要求96所述的方法,其中,该细胞是位于受试者内。

98.如权利要求97所述的方法,其中,该受试者是人。

99.如权利要求98所述的方法,其中,该受试者患有C9orf72相关疾患。

100.如权利要求99所述的方法,其中,该C9orf72相关疾患选自由C9orf72肌萎缩侧索硬化/额颞叶痴呆或由于C9orf72扩增引起的亨廷顿样综合征所组成的组。

101.如权利要求96至100中任一项所述的方法,其中,使该细胞与该dsRNA剂接触将C9orf72的表达抑制不超过50%。

102.如权利要求96至101中任一项所述的方法,其中,抑制C9orf72的表达将该受试者

血清中的C9orf72蛋白质水平降低不超过50％。

103.一种治疗患有将受益于C9orf72表达减少的疾患的受试者的方法，该方法包括对该受试者给药治疗有效量的如权利要求1至87中任一项所述的dsRNA剂或如权利要求89至95中任一项所述的药物组合物，从而治疗患有将受益于C9orf72表达减少的疾患的受试者。

104.一种预防患有将受益于C9orf72表达减少的疾患的受试者的至少一种症状的方法，该方法包括对该受试者给药预防有效量的如权利要求1至87中任一项所述的dsRNA剂或如权利要求89至95中任一项所述的药物组合物，从而预防患有将受益于C9orf72表达减少的疾患的受试者的至少一种症状。

105.如权利要求103或104所述的方法，其中，该疾患是C9orf72相关疾患。

106.如权利要求105所述的方法，其中，该C9orf72相关疾患选自由C9orf72肌萎缩侧索硬化/额颞叶痴呆及由于C9orf72扩增引起的亨廷顿样综合征所组成的组。

107.如权利要求106所述的方法，其中，该受试者是人。

108.如权利要求106或107所述的方法，其中，将该剂给药至该受试者造成RNA团簇的重复序列长度依赖性形成，特异性RNA结合蛋白质的隔离，或者聚(甘氨酸-丙氨酸)肽、聚(甘氨酸-脯氨酸)肽、聚(甘氨酸-精氨酸)肽、聚(丙氨酸-脯氨酸)肽或聚(脯氨酸-精氨酸)肽的累积或集聚的降低。

109.如权利要求108的所述的方法，其中，该RNA团簇的重复序列长度依赖性形成的降低，特异性RNA结合蛋白质的隔离，或者聚(甘氨酸-丙氨酸)肽、聚(甘氨酸-脯氨酸)肽、聚(甘氨酸-精氨酸)肽、聚(丙氨酸-脯氨酸)肽或聚(脯氨酸-精氨酸)肽的累积或集聚被降低不超过50％；并且，该

C9orf72成熟RNA的表达被抑制少于50％。

110.如权利要求103至109中任一项所述的方法，其中，该dsRNA剂是以约0.01mg/kg至约50mg/kg的剂量给药至该受试者。

111.如权利要求103至110中任一项所述的方法，其中，该dsRNA剂是经皮下给药至该受试者。

112.如权利要求93至110中任一项所述的方法，其中，该dsRNA剂是经鞘内给药至该受试者。

113.如权利要求93至112中任一项所述的方法，进一步包括测定来自该受试者的样本中C9orf72的水平。

114.如权利要求113所述的方法，其中，该受试者样本中C9orf72的水平是血液、血清或脑脊液样本中的C9orf72蛋白质水平。

115.如权利要求103至114中任一项所述的方法，进一步包括将附加治疗剂给药至该受试者。

116.一种试剂盒，其包含如权利要求1至87中任一项所述的dsRNA剂或如权利要求89至95中任一项所述的药物组合物。

117.一种小瓶，其包含如权利要求1至87中任一项所述的dsRNA剂或如权利要求89至95中任一项所述的药物组合物。

118.一种注射器，其包含如权利要求1至87中任一项所述的dsRNA剂或如权利要求89至95中任一项所述的药物组合物。

人类染色体9开放阅读框72(C9ORF72)iRNA剂组合物及其使用方法

[0001] 相关申请

[0002] 本申请要求2019年12月13日递交的美国临时申请第62/947,605号及2019年12月13日递交的美国临时申请第62/947,768号的优先权权益。前述申请的全部内容通过援引并入本文。

[0003] 序列表

[0004] 本申请含有序列表,其已经以ASCII格式提交电子版,且通过援引以其整体并入本文。所述ASCII副本于2020年XXX创建,名为121301_10220_SL.txt,大小为XXX字节。

背景技术

[0005] 人类染色体9开放阅读框72(C9orf72)是由c9orf72基因编码的蛋白质。C9orf72见于脑部多个区域诸如大脑皮质内、神经元的细胞质内以及突触前末端内。

[0006] 转录可选起始及终止位点的分化用途从C9orf72 DNA生成三种RNA转录本。这些编码由衍生自变体2(NM_018325.4)及3(NM_001256054.2)的大约54kDa的长同种型(同种型A)与衍生自变体1(NM_145005.6)的大约24kDa的短同种型(同种型B)组成的两种蛋白质同种型(参见,例如,Barker,et al.(2017)Frontiers Cell Neurosci 11:1-15的图1)。

[0007] C9orf72基因的两个可选使用的第一外显子是外显子1a及1b(参见,例如,上述Barker,et al.,的图1)。已经表明,C9orf72基因的位于外显子1a与1b之间的第一内含子中的大GGGGCC(G_4C_2)六核苷酸重复扩增(从约2-22个拷贝到700-1600个)干扰C9orf72蛋白质的功能并且具有致病性,其导致几种具有截然不同的临床特征但具有共享的病理学特征及遗传成因的神经退行性疾病(Ling,et al.(2013)Neuron 79:416-438)。特别地,C9orf72基因中六核苷酸重复扩增的存在是家族性及散发性肌萎缩侧索硬化(Amyotrophic Lateral Sclerosis,ALS)的最常见遗传成因,该疾病是致命性的脑及脊髓中运动神经元的退行性疾病。事实上,C9orf72突变六核苷酸重复扩增存在于大于40%的家族性ALS受试者中及8%至10%的散发性ALS受试者中。C9orf72基因中的六核苷酸重复扩增也是额颞叶痴呆(Frontotemporal Dementia,FTD)的最常见的家族性成因,FTD是阿兹海默氏症后的老年前期痴呆的第二常见形式,其以行为及语言缺陷为特征并且在病理上表达为脑中额叶及前颞叶的神经元萎缩。由于C9orf72扩增引起的亨廷顿样综合征(Huntington-Like Syndrome Due To C9orf72Expansions),以运动障碍(包括张力失常、舞蹈症、肌阵挛、震颤及强直)、认知及记忆损害、早期精神紊乱及行为问题为特征,也与C9orf72基因中的六核苷酸重复扩增相关。

[0008] 尽管C9orf72蛋白质的功能尚在探究之中,但已经表明,C9orf72与Rab蛋白质交互作用并将其激活,Rab蛋白质参与调控细胞骨架、自噬作用及胞吞运输。此外,已经证明,在与C9orf72六核苷酸重复扩增相关的神经退行性疾病中,大量细胞途径失调。例如,改变的RNA加工一直是C9orf72疾病研究的前沿。这包括重复序列的双向转录,重复RNA累积到核团簇(foci)处从而隔离特异性RNA结合蛋白(RBP),以及借由重复相关的非AUG(RAN)启动的翻

译将RNA重复翻译为二肽重复序列蛋白(DPR)。此外,已经表明,C9orf72 RNA从RNA聚合酶II中的释放、于细胞质中的翻译及降解的中断是被C9orf72六核苷酸重复扩增所中断的。此外,就C9orf72 RNA的转录、剪接及定位而言,在其自身的加工中已经鉴定出若干改变(参见,例如,上述Barker,et al.)。

[0009] 无论机制如何,几个小组已经鉴定出,在患有C9orf72相关疾病的受试者神经系统的几种细胞类型中,含正义及反义C9orf72团簇的存在以及异常二肽重复(DPR)蛋白质(poly(GA)、poly(GR)、poly(GP)、poly(PA)及poly(PR))的存在,这些是通过重复相关的非AUG依赖性(RAN)翻译从含正义或反义重复的C9orf72 RNA的所有阅读框产生(Lagier-Tourenne,et al.(2013)Proc Natl Acad Sci USA doi/10.1073/pnas.1318835110;Jiang,et al.(2016)Neuron 90:535-550)。此外,在具有一个未激活的C9orf72等位基因的小鼠中,未引起疾病,但在具有两个未激活的C9orf72等位基因的小鼠中,观察到脾肿大、淋巴结增大及轻度社交缺陷,但未观察到运动功能异常。此外,在表达具有至多450个GGGGCC重复的人类C9orf72 RNA的小鼠中显示,六核苷酸扩增造成通过AUG非依赖性翻译合成的二肽重复蛋白质以及含有正义及反义RNA的团簇的依赖于年龄、重复序列长度及表达水平的累积,伴有海马神经元损失、焦虑增加以及认知功能受损(Jiang,et al.(2016)Neuron 90:535-550)。

[0010] 患有C9orf72相关疾病(例如,C9orf72肌萎缩侧索硬化、C9orf72额颞叶痴呆或亨廷顿氏病,例如,由于C9orf72扩增引起的亨廷顿样综合征)的受试者目前尚无治愈法,并且治疗的目的仅在于随着疾病进展而缓解症状及改善患者生命质量。

[0011] 因此,本领域需要可以选择性且有效地抑制C9orf72基因表达的药剂,用于例如治疗患有C9orf72相关疾患的受试者。

发明内容

[0012] 本公开提供RNAi组合物,其影响C9orf72基因(诸如具有经扩增的GGGGCC(G_4C_2)重复的C9orf72基因)的RNA转录本的RNA诱导沉默复合物(RISC)介导的裂解。C9orf72基因可处于细胞内,例如,受试者诸如人体内的细胞内。这些iRNA的使用使得哺乳动物体内相应基因(C9orf72基因)的mRNA的靶向降解成为可能。

[0013] 本发明的iRNA已经设计为靶向C9orf72基因,例如,在该基因的内含子中具有经扩增的GGGGCC六核苷酸重复并且具有核苷酸修饰的组合的C9orf72基因。这些剂可靶向成熟C9orf72 mRNA(其内含子被剪除的mRNA)或含有六核苷酸重复的C9orf72 RNA(含有C9orf72内含子1A的RNA)。本发明的iRNA可降低C9orf72成熟mRNA的水平,降低的水平小于其降低含有六核苷酸重复的C9orf72RNA的水平。例如,本发明的iRNA可将C9orf72成熟mRNA的水平降低不超过约50%,并且减少含正义及反义的团簇及异常二肽重复(DPR)蛋白质(poly(GA)、poly(GR)、poly(GP)、poly(PA)及poly(PR))的水平,和/或将含有六核苷酸重复的C9orf72 RNA水平降低超过约50%,并且减少含正义及反义的团簇及异常二肽重复(DPR)蛋白质(poly(GA)、poly(GR)、poly(GP)、poly(PA)及poly(PR))的水平。不试图受限于理论,据信,前述特性与这些iRNA中的特定靶标位点或特定修饰的组合或亚组合赋予本发明的iRNA以改善的功效、稳定性、效力、耐久性及安全性。

[0014] 一方面,本发明提供一种双链核糖核酸(dsRNA)剂,用于敲低细胞中的C9orf72靶

标RNA。

[0015] 在一个实施方案中,dsRNA剂靶向含有六核苷酸重复的C9orf72靶标RNA区域。

[0016] 一方面,本发明提供双链核糖核酸(dsRNA)剂,其用于抑制C9orf72(诸如含有六核苷酸重复的C9orf72 RNA)的表达,其中,该dsRNA剂包含形成双链区域的正义链及反义链,其中,该反义链包含与SEQ ID NO:125的核苷酸序列相差不超过3个核苷酸的至少15个连续核苷酸,以及,其中,该正义链、该反义链、或该正义链及反义链两者缀合至一个或多个亲脂性部分。

[0017] 另一方面,本发明提供双链核糖核酸(dsRNA)剂,其用于抑制C9orf72(诸如含有六核苷酸重复的C9orf72 RNA)的表达,其中,该dsRNA剂包含形成双链区域的正义链及反义链,其中,该正义链包含与SEQ ID NO:133的核苷酸200至290相差不超过三个核苷酸的至少15个连续核苷酸,以及该反义链包含来自SEQ ID NO:134的相应核苷酸序列的至少15个连续核苷酸,以及,其中,该正义链、该反义链、或该正义链及反义链二者缀合至一个或多个亲脂性部分。

[0018] 一方面,本发明提供双链核糖核酸(dsRNA)剂,其用于抑制C9orf72(诸如含有六核苷酸重复的C9orf72 RNA)的表达,其中,该dsRNA剂包含形成双链区域的正义链及反义链,其中,该正义链包含与表10A、10B、12至15、19、20或21任一者中的反义核苷酸序列的任一个相差不超过3个核苷酸的至少15个连续核苷酸

[0019] 在一个实施方案中,该正义链及反义链的核苷酸序列包含表12或13任一者中的正义链及反义链核苷酸序列的任一个。在一个实施方案中,该反义链的核苷酸序列包含表12或13任一者中的反义链核苷酸序列的任一个。

[0020] 在一个实施方案中,该正义链包含与SEQ ID NO:121的核苷酸230至270、233至262、800至840、800至830、802至828、1240至1290、1240至1280、1247至1288、1590至1645、1590至1620、及1594至1642中任一核苷酸序列相差不超过三个核苷酸的至少15个连续核苷酸,并且该反义链包含来自SEQ ID NO:125的相应核苷酸序列的至少15个连续核苷酸。

[0021] 在一个实施方案中,该正义链包含与SEQ ID NO:121的核苷酸1594至1616、802至824、239至261、1308至1330、233至255、1595至1617、240至262、1532至1554、237至259、3268至3290、806至828、1620至1642、526至548、1169至1191、1266至1288、1247至1269、586至608、1257至1279、及400至422中任一核苷酸序列相差不超过三个核苷酸的至少15个连续核苷酸,并且该反义链包含来自SEQ ID NO:125的相应核苷酸序列的至少15个连续核苷酸。

[0022] 在一个实施方案中,该反义链包含与选自由下列所组成的组的双链体的反义链核苷酸序列的任一个相差不超过三个核苷酸的至少15个连续核苷酸:AD-348904.1、AD-348136.1、AD-347612.1、AD-348639.1、AD-347606.1、AD-348905.1、AD-347613.1、AD-348842.1、AD-347610.1、AD-350329.1、AD-348140.1、AD-348930.1、AD-347863.1、AD-348500.1、AD-348597.1、AD-348578.1、AD-347923.1、AD-348588.1、及AD-347773.1。

[0023] 在一个实施方案中,该正义链及反义链的核苷酸序列包含表14、15、19、20或21任一者中的正义链及反义链核苷酸序列的任一个。

[0024] 一方面,本发明提供一种用于选择性地抑制包含六核苷酸重复的C9orf72的表达的双链核糖核酸(dsRNA)剂,该重复包含SEQ ID NO:1的多个连续拷贝,其中,该dsRNA剂包含形成双链区域的正义链及反义链,其中,该反义链包含与表14、15或21任一者中的反义核

苷酸序列的任一个相差不超过3个核苷酸的至少15个连续核苷酸。

[0025] 在一个实施方案中,该反义链包含与选自由下列所组成的组的双链体的反义链核苷酸序列的任一个相差不超过三个核苷酸的至少15个连续核苷酸:AD-463863.1、AD-463862.1、AD-463869.1、AD-463873.1、AD-463872.1及AD-463860.1。

[0026] 在一个实施方案中,该反义链包含与选自由下列所组成的组的双链体的反义链核苷酸序列的任一个相差不超过三个核苷酸的至少15、16、17、18、19或20个连续核苷酸:AD-463863.1、AD-463862.1、AD-463869.1、AD-463873.1、AD-463872.1及AD-463860.1。在一些实施方案中,该正义及反义链各自独立为19、20、21、22、23、24或25个核苷酸的长度。

[0027] 在一个实施方案中,该反义链包含与选自由下列所组成的组的双链体的反义链核苷酸序列的任一个相差不超过两个核苷酸的至少15、16、17、18、19或20个连续核苷酸:AD-463863.1、AD-463862.1、AD-463869.1、AD-463873.1、AD-463872.1及AD-463860.1。在一些实施方案中,该正义及反义链各自独立为19、20、21、22、23、24或25个核苷酸的长度。

[0028] 在一个实施方案中,该反义链包含与选自由下列所组成的组的双链体的反义链核苷酸序列的任一个相差不超过一个核苷酸的至少15、16、17、18、19或20个连续核苷酸:AD-463863.1、AD-463862.1、AD-463869.1、AD-463873.1、AD-463872.1及AD-463860.1。在一些实施方案中,该正义及反义链各自独立为19、20、21、22、23、24或25个核苷酸的长度。

[0029] 在一个实施方案中,该反义链包含来自选自由下列所组成的组的双链体的反义链核苷酸序列的任一个的至少15、16、17、18、19或20个连续核苷酸:AD-463863.1、AD-463862.1、AD-463869.1、AD-463873.1、AD-463872.1及AD-463860.1。在一些实施方案中,该正义及反义链各自独立为19、20、21、22、23、24或25个核苷酸的长度。

[0030] 在一个实施方案中,该C9orf72靶标RNA包含含有SEQ ID NO:1的多个连续拷贝的六核苷酸重复,该dsRNA剂包含正义链及反义链,该反义链包含与C9orf72靶标RNA互补的区域,并且该互补的区域包含与位于C9orf72靶标RNA的外显子1A起始位点与外显子2起始位点之间的核苷酸序列的反向互补序列相差不超过3个核苷酸的至少15个连续核苷酸。

[0031] 在一个实施方案中,该C9orf72靶标RNA包含含有SEQ ID NO:1的多个连续拷贝的六核苷酸重复,该dsRNA剂包含正义链及反义链,并且该反义链包含与存在于C9orf72靶标RNA的外显子1A起始位点与外显子2起始位点之间的序列中的15至25个连续核苷酸序列具有至少80%互补性的15至25个连续核苷酸序列。

[0032] 在一个实施方案中,该互补的区域包含与位于C9orf72靶标RNA的外显子1A与外显子1B之间的核苷酸序列相差不超过3个核苷酸的至少15个连续核苷酸的反向互补序列。

[0033] 在一个实施方案中,该互补的区域包含与位于C9orf72靶标RNA的外显子1A与六核苷酸重复之间的核苷酸序列相差不超过3个核苷酸的至少15个连续核苷酸的反向互补序列。

[0034] 在一个实施方案中,该正义链包含与SEQ ID NO:115的核苷酸序列相差不超过3个核苷酸的至少15个连续核苷酸,且该反义链包含与SEQ ID NO:116的核苷酸序列相差不超过3个核苷酸的至少15个连续核苷酸。

[0035] 在一个实施方案中,该正义链包含与SEQ ID NO:117的核苷酸序列相差不超过3个核苷酸的至少15个连续核苷酸,且该反义链包含与SEQ ID NO:118的核苷酸序列相差不超过3个核苷酸的至少15个连续核苷酸。

[0036] 在一个实施方案中，该反义链包含与SEQ ID NO:3、5、7、9、11及13中反义链序列的任一个相差不超过3个核苷酸的至少15个连续核苷酸，和/或该正义链包含与SEQ ID NO:2、4、6、8、10及12中正义链序列的任一个相差不超过3个核苷酸的至少15个连续核苷酸。

[0037] 在一个实施方案中，位于C9orf72靶标RNA的外显子1A起始位点与外显子2起始位点之间的核苷酸序列不存在于成熟的C9orf72信使RNA中。

[0038] 在一个实施方案中，该正义链、该反义链、或该正义链及该反义链二者缀合至一个或多个亲脂性部分。

[0039] 在一个实施方案中，该亲脂性部分缀合至该dsRNA剂的该双链区域中的一个或多个内部位置。

[0040] 在一个实施方案中，亲脂性部分经由接头或载剂缀合。

[0041] 在一个实施方案中，通过logKow测量，该亲脂性部分的亲脂性是超过0。

[0042] 在一个实施方案中，通过该双链RNAi剂的血浆蛋白结合测定中的未结合部分测量，该双链RNAi剂的疏水性是超过0.2。

[0043] 在一个实施方案中，该血浆蛋白结合测定是使用人血清白蛋白蛋白质的电泳迁移率变动分析。

[0044] 在一些实施方案中，该dsRNA剂包含至少一个经修饰的核苷酸。

[0045] 在一个实施方案中，该正义链的不超过五个核苷酸及该反义链的不超过五个核苷酸是未经修饰的核苷酸。

[0046] 在一个实施方案中，该正义链的全部核苷酸及该反义链的全部核苷酸是经修饰的核苷酸。

[0047] 在一个实施方案中，该经修饰的核苷酸的至少一个选自由下列所组成的组：脱氧核苷酸、3'-端脱氧胸腺嘧啶(dT)核苷酸、2'-O-甲基修饰的核苷酸、2'-氟修饰的核苷酸、2'-脱氧修饰的核苷酸、锁核苷酸、未锁核苷酸、构象限制性核苷酸、约束的乙基核苷酸、无碱基的核苷酸、2'-氨基修饰的核苷酸、2'-O-烯丙基修饰的核苷酸、2'-C-烷基修饰的核苷酸、2'-羟基修饰的核苷酸、2'-甲氧基乙基修饰的核苷酸、2'-O-烷基修饰的核苷酸、吗啉代核苷酸、氨基磷酸酯、包含非天然碱基的核苷酸、四氢呋喃修饰的核苷酸、1,5-失水己糖醇修饰的核苷酸、环己基修饰的核苷酸、包含5'-硫代磷酸酯基团的核苷酸、包含5'-甲基膦酸酯基团的核苷酸、包含5'-磷酸酯或5'-磷酸酯模拟物的核苷酸、包含磷酸乙烯酯的核苷酸、包含腺苷-二醇核酸(GNA)的核苷酸、包含胸苷二醇核酸(GNA)S异构体的核苷酸、包含2-羟甲基-四氢呋喃-5-磷酸酯的核苷酸、包含2'-脱氧胸苷-3'磷酸酯的核苷酸、包含2'-脱氧鸟苷-3'磷酸酯的核苷酸、连接至胆固醇基衍生物的末端核苷酸、及十二酸双癸基酰胺基团；及其组合。

[0048] 在一个实施方案中，该经修饰的核苷酸选自由下列所组成的组：2'-脱氧-2'-氟修饰的核苷酸、2'-脱氧修饰的核苷酸、3'-末端脱氧-胸腺嘧啶核苷酸(dT)、锁核苷酸、无碱基的核苷酸、2'-氨基修饰的核苷酸、2'-烷基修饰的核苷酸、吗啉代核苷酸、氨基磷酸酯、以及包含非天然碱基的核苷酸。

[0049] 在一个实施方案中，该经修饰的核苷酸包含3'-末端脱氧-胸腺嘧啶核苷酸(dT)的短序列。

[0050] 在一个实施方案中，核苷酸的修饰是2'-O-甲基、GNA及2'-氟修饰。

[0051] 在一个实施方案中，该正义链的实质上全部核苷酸包含选自由2'-O-甲基修饰及2'-氟修饰所组成的组的修饰。在一个实施方案中，该正义链的全部核苷酸包含选自由2'-O-甲基修饰及2'-氟修饰所组成的组的修饰。

[0052] 在一个实施方案中，该反义链的实质上全部核苷酸包含选自由2'-O-甲基修饰及2'-氟修饰所组成的组的修饰。在一些实施方案中，该反义链的全部核苷酸包含选自由2'-O-甲基修饰及2'-氟修饰所组成的组的修饰。

[0053] 在一个实施方案中，该正义链的实质上全部核苷酸包含选自由2'-O-甲基修饰、2'-氟修饰及二醇核酸(GNA)修饰所组成的组的修饰。在一些实施方案中，该正义链的全部核苷酸包含选自由2'-O-甲基修饰、2'-氟修饰及二醇核酸(GNA)修饰所组成的组的修饰。

[0054] 在一个实施方案中，该反义链的实质上全部核苷酸包含选自由2'-O-甲基修饰、2'-氟修饰及二醇核酸(GNA)修饰所组成的组的修饰。在一些实施方案中，该反义链的全部核苷酸包含选自由2'-O-甲基修饰、2'-氟修饰及二醇核酸(GNA)修饰所组成的组的修饰。

[0055] 在一些实施方案中，该dsRNA剂进一步包含至少一个硫代磷酸酯核苷酸间键。

[0056] 在一个实施方案中，该dsRNA剂进一步包含6至8个硫代磷酸酯核苷酸间键。

[0057] 在一个实施方案中，该正义链包含至少一个硫代磷酸酯或甲基膦酸酯核苷酸间键，并且该反义链包含至少一个硫代磷酸酯或甲基膦酸酯核苷酸间键。

[0058] 在一个实施方案中，该正义链包含至少两个硫代磷酸酯或甲基膦酸酯核苷酸间键。

[0059] 在一个实施方案中，该反义链包含至少两个、至少三个或至少四个硫代磷酸酯或甲基膦酸酯核苷酸间键。

[0060] 在一个实施方案中，该至少一个硫代磷酸酯或甲基膦酸酯核苷酸间键位于一链的5'末端、位于一链的3'末端、或位于一链的5'末端及3'末端两处。

[0061] 在一个实施方案中，该至少一个硫代磷酸酯或甲基膦酸酯核苷酸间键位于正义链的5'-端。在一个实施方案中，该正义链包含两个位于5'-端的硫代磷酸酯核苷酸间键。

[0062] 在一个实施方案中，该至少一个硫代磷酸酯或甲基膦酸酯核苷酸间键位于反义链的5'-端及3'-端两处。在一些实施方案中，该反义链包含两个位于5'-端的硫代磷酸酯核苷酸间键以及两个位于3'-端的硫代磷酸酯核苷酸间键。

[0063] 在一个实施方案中，该正义链的全部核苷酸包含选自由2'-O-甲基修饰及2'-氟修饰所组成的组的修饰，该反义链的全部核苷酸包含选自由2'-O-甲基修饰及2'-氟修饰所组成的组的修饰，该正义链包含位于5'端的两个硫代磷酸酯核苷酸间键，并且该反义链包含位于5'端的两个硫代磷酸酯核苷酸间键以及位于3'端的两个硫代磷酸酯核苷酸间键。

[0064] 在一个实施方案中，(I)该正义链包含：(a)位于多个核苷酸处的2'-O-甲基修饰；(b)位于多个核苷酸处的2'-氟修饰；以及(c)多个硫代磷酸酯核苷酸间键；并且，(II)该反义链包含：(a)位于多个核苷酸处的2'-O-甲基修饰；(c)位于多个核苷酸处的2'-氟修饰；以及(d)位于多个核苷酸之间的硫代磷酸酯核苷酸间键。任选地，该dsRNA剂包含位于该反义链的3'末端的突出端，以及位于该反义链的5'末端的钝端。

[0065] 在一个实施方案中，(I)该正义链包含：(a)21个核苷酸的长度；(b)位于自5'末端计数的核苷酸位置1至6、8及12至21处的2'-O-甲基修饰；(c)位于自5'末端计数的核苷酸位置7及9至11处的2'-氟修饰；以及(d)位于自5'末端计数的核苷酸位置1与2之间及核苷酸位

置2与3之间的硫代磷酸酯核苷酸间键;并且,(II)该反义链包含:(a)23个核苷酸的长度;(b)位于自5'末端计数的核苷酸位置1、3至5、7、10至13、15及17至23处的2'-O-甲基修饰;(c)位于自5'末端计数的核苷酸位置2、6、8、9、14及16处的2'-氟修饰;以及(d)位于(自5'末端计数)核苷酸位置1与2之间、核苷酸位置2与3之间、核苷酸位置21与22之间及核苷酸位置22与23之间的硫代磷酸酯核苷酸间键,其中,该dsRNA剂包含位于反义链3'末端的两个核苷酸的突出端以及位于反义链5'末端的钝端。

[0066] 在一个实施方案中,该反义链包含SEQ ID NO:3、5、7、9、11及13中任一反义链序列,和/或其中该正义链包含SEQ ID NO:2、4、6、8、10及12中任一正义链序列。

[0067] 在一个实施方案中,该正义链是不超过30个核苷酸的长度。在另一个实施方案中,该反义链是不超过30个核苷酸的长度。在一个实施方案中,该正义链及反义链各自独立地为不超过30个核苷酸的长度。

[0068] 在一个实施方案中,至少一链包含至少1个核苷酸的3'突出端。在另一个实施方案中,至少一链包含至少2个核苷酸的3'突出端。在一个实施方案中,该反义链包含3'-突出端。

[0069] 双链区域可是15至30个核苷酸对的长度;17至23个核苷酸对的长度;17至25个核苷酸对的长度;23至27个核苷酸对的长度;19至21个核苷酸对的长度;21至23个核苷酸对的长度;或21个核苷酸的长度。

[0070] 该正义链及该反义链可各自独立地为19至30个核苷酸的长度;19至25个核苷酸对的长度;19至23个核苷酸对的长度;或21至23个核苷酸对的长度;或21个核苷酸的长度。

[0071] 在一个实施方案中,该互补区域是至少17个核苷酸的长度。在其他实施方案中,该互补区域是19至30个核苷酸的长度,19至25个核苷酸的长度,或21至23个核苷酸的长度。

[0072] 在一个实施方案中,该互补区域是至少17个核苷酸的长度。在其他实施方案中,该互补区域是19至30个核苷酸的长度,19至25个核苷酸的长度,或21至23个核苷酸的长度。

[0073] 在一个实施方案中,该互补区域与位于C9orf72基因的外显子1A起始位点与外显子2起始位点之间的序列具有至少85%的互补性。在一些实施方案中,该反义链包含15至25个连续核苷酸的序列,该序列与存在于C9orf72靶标RNA的外显子1A起始位点与外显子2起始位点之间的序列中的15至25个连续核苷酸序列具有至少85%的互补性。在其他实施方案中,该互补区域与位于C9orf72靶标RNA的外显子1A起始位点与外显子2起始位点之间的序列具有至少90%的互补性。在一个实施方案中,该互补区域与位于C9orf72靶标RNA的外显子1A起始位点与外显子2起始位点之间的序列具有至少95%的互补性。在一些实施方案中,该互补区域与位于C9orf72靶标RNA的外显子1A起始位点与外显子2起始位点之间的序列具有100%的互补性。在一些实施方案中,该互补区域与位于C9orf72靶标RNA的外显子1A末端与六核苷酸重复区域起始位点之间的序列具有100%的互补性。

[0074] 在一个实施方案中,该互补区域与位于C9orf72基因的外显子1A末端与内含子1A中的六核苷酸重复起始位点之间的序列具有至少85%的互补性。在一些实施方案中,该反义链包含15至25个连续核苷酸的序列,该序列与存在于C9orf72靶标RNA的外显子1A末端与内含子1A中的六核苷酸重复起始位点之间的序列中的15至25个连续核苷酸序列具有至少85%的互补性。在其他实施方案中,该互补区域与位于C9orf72靶标RNA的外显子1A末端与内含子1A中的六核苷酸重复起始位点之间的序列具有至少90%的互补性。在一个实施方案

中，该互补区域与位于C9orf72靶标RNA的外显子1A末端与内含子1A中的六核苷酸重复起始位点之间的序列具有至少95％的互补性。在一些实施方案中，该互补区域与位于C9orf72靶标RNA的外显子1A末端与内含子1A中的六核苷酸重复起始位点之间的序列具有100％的互补性。

[0075] 在一个实施方案中，该一个或多个亲脂性部分经由诸如接头或载剂缀合至至少一链的一个或多个内部位置。

[0076] 在一个实施方案中，该内部位置包括来自至少一链的每一端的末端两个位置以外的所有位置。

[0077] 在另一个实施方案中，该内部位置包括来自至少一链的每一端的末端三个位置以外的所有位置。

[0078] 在一个实施方案中，这些内部位置排除正义链的裂解位点区域。

[0079] 在一个实施方案中，该内部位置包括从该正义链的5’末端起计数的位置9至12以外的所有位置。

[0080] 在另一个实施方案中，该内部位置包括从该正义链的3’末端起计数的位置11至13以外的所有位置。

[0081] 在一个实施方案中，这些内部位置排除反义链的裂解位点区域。

[0082] 在一个实施方案中，该内部位置包括从该反义链的5’-端起计数的位置12至14以外的所有位置。

[0083] 在一个实施方案中，该内部位置包括该正义链从3’-端起计数的位置11至13以及该反义链从5’-端起计数的位置12至14以外的所有位置。

[0084] 在一个实施方案中，该一个或多个亲脂性部分缀合至一个或多个选自由下列所组成的组的内部位置：该正义链的位置4至8及13至18，以及该反义链上的位置6至10及15至18，从每一链5’端起计数。

[0085] 在另一个实施方案中，该一个或多个亲脂性部分缀合至一个或多个选自由下列所组成的组的内部位置：该正义链的位置5、6、7、15及17，以及该反义链上的位置15及17，从每一链5’端起计数。

[0086] 在一个实施方案中，该双链区域中的内部位置排除该正义链的裂解位点区域。

[0087] 在一个实施方案中，该正义链是21个核苷酸的长度，该反义链是23个核苷酸的长度，以及，该亲脂性部分缀合至该正义链的位置21、位置20、位置15、位置1、位置7、位置6或位置2或该反义链的位置16。

[0088] 在一个实施方案中，该亲脂性部分缀合至该正义链的位置21、位置20、位置15、位置1或位置7。

[0089] 在另一个实施方案中，该亲脂性部分缀合至该正义链的位置21、位置20或位置15。

[0090] 在又一个实施方案中，该亲脂性部分缀合至该正义链的位置20或位置15。

[0091] 在一个实施方案中，该亲脂性部分缀合至该反义链的位置16。

[0092] 在一个实施方案中，亲脂性部分是脂族、脂环族或多脂环族化合物。

[0093] 在一个实施方案中，该亲脂性部分选自由下列所组成的组：脂质、胆固醇、视黄酸、胆酸、金刚烷乙酸、1-芘丁酸、双氢睾酮、1,3-双-O(十六烷基)甘油、香叶基氧己醇、十六烷基甘油、冰片、薄荷醇、1,3-丙二醇、十七烷基基团、棕榈酸、肉豆蔻酸、03-(油酰基)石胆酸、

O3-(油酰基)胆烯酸、二甲氧基三苯甲基或吩噁嗪(phenoxazine)。

[0094] 在一个实施方案中,该亲脂性部分含有饱和或不饱和的C4-C30烃链,以及任选的选自由羟基、胺、羧酸、磺酸酯、磷酸酯、硫醇、叠氮化物及炔所组成的组的官能基团。

[0095] 在一个实施方案中,该亲脂性部分含有饱和或不饱和的C6-C18烃链。

[0096] 在一个实施方案中,该亲脂性部分含有饱和或不饱和的C16烃链。

[0097] 在一个实施方案中,该饱和或不饱和的C16烃链缀合至自该链的5'-端起计数的位置6。

[0098] 在一个实施方案中,该亲脂性部分经由载剂缀合,该载剂替换该内部位置或该双链区域中的一个或多个核苷酸。

[0099] 在一个实施方案中,该载剂是选自由下列所组成的组的环状基团:吡咯烷基、吡唑啉基、吡唑烷基、咪唑啉基、咪唑烷基、哌啶基、哌嗪基(piperazinyl)、[1,3]二氧杂环戊基、噁唑烷基、异噁唑烷基、吗啉基、噻唑烷基、异噻唑烷基、喹喔啉基、哒嗪酮基(pyridazinonyl)、四氢呋喃基及十氢萘基;或是基于丝氨醇主链或二乙醇胺主链的非环状部分。

[0100] 在一个实施方案中,该亲脂性部分经由接头缀合至该双链iRNA剂,该接头含有醚、硫醚、脲、碳酸酯、胺、酰胺、马来酰亚胺-硫醚、二硫化物、磷酸二酯、磺酰胺键、点击反应的产物或氨基甲酸酯。

[0101] 在一个实施方案中,亲脂性部分缀合至核碱基、糖部分或核苷酸间键。

[0102] 在一个实施方案中,该亲脂性部分或靶向配体经由生物可裂解的接头缀合,该接头是选自由下列所组成的组:DAN,RNA,二硫化物,酰胺,半乳糖胺、葡糖胺、葡萄糖、半乳糖、甘露糖的官能化的单糖或寡糖,及其组合。

[0103] 在一个实施方案中,该正义链的3'端经由端帽保护,该端帽具有胺的环状基团,所述环状基团选自由下列所组成的组:吡咯烷基、吡唑啉基、吡唑烷基、咪唑啉基、咪唑烷基、哌啶基、哌嗪基、[1,3]二氧杂环戊基、噁唑烷基、异噁唑烷基、吗啉基、噻唑烷基、异噻唑烷基、喹喔啉基、哒嗪酮基、四氢呋喃基及十氢萘基。

[0104] 在一个实施方案中,该dsRNA剂进一步包含靶向肝组织的靶向配体。

[0105] 在一个实施方案中,该靶向配体是GalNAc缀合物。

[0106] 在一个实施方案中,该dsRNA剂进一步包含末端手性修饰,其出现于该反义链3'端的第一个核苷酸间键处,具有Sp构型的键磷原子;末端手性修饰,其出现于该反义链5'端的第一个核苷酸间键处,具有Rp构型的键磷原子;以及末端手性修饰,其出现于该正义链5'端的第一个核苷酸间键处,具有Rp构型或Sp构型的键磷原子。

[0107] 在另一个实施方案中,该dsRNA剂进一步包含末端手性修饰,其出现于该反义链3'端的第一及第二个核苷酸间键处,具有Sp构型的键磷原子;末端手性修饰,其出现于该反义链5'端的第一个核苷酸间键处,具有Rp构型的键磷原子;以及末端手性修饰,其出现于该正义链5'端的第一个核苷酸间键处,具有Rp构型或Sp构型的键磷原子。

[0108] 在又一个实施方案中,该dsRNA剂进一步包含末端手性修饰,其出现于该反义链3'端的第一、第二及第三个核苷酸间键处,具有Sp构型的键磷原子;末端手性修饰,其出现于该反义链5'端的第一个核苷酸间键处,具有Rp构型的键磷原子;以及末端手性修饰,其出现于该正义链5'端的第一个核苷酸间键处,具有Rp构型或Sp构型的键磷原子。

[0109] 在另一个实施方案中，该dsRNA剂进一步包含末端手性修饰，其出现于该反义链3’端的第一及第二个核苷酸间键处，具有Sp构型的键磷原子；末端手性修饰，其出现于该反义链3’端的第三个核苷酸间键处，具有Rp构型的键磷原子；末端手性修饰，其出现于该反义链5’端的第一个核苷酸间键处，具有Rp构型的键磷原子；以及末端手性修饰，其出现于该正义链5’端的第一个核苷酸间键处，具有Rp构型或Sp构型的键磷原子。

[0110] 在另一个实施方案中，该dsRNA剂进一步包含末端手性修饰，其出现于该反义链3’端的第一及第二个核苷酸间键处，具有Sp构型的键磷原子；末端手性修饰，其出现于该反义链5’端的第一及第二个核苷酸间键处，具有Rp构型的键磷原子；以及末端手性修饰，其出现于该正义链5’端的第一个核苷酸间键处，具有Rp构型或Sp构型的键磷原子。

[0111] 在一个实施方案中，该dsRNA剂包含位于该反义链的5’端的磷酸酯或磷酸酯模拟物。

[0112] 在一个实施方案中，该磷酸酯模拟物是5’-乙烯基磷酸酯(VP)。

[0113] 在一个实施方案中，位于该双链体反义链的5’端第1位置处的碱基对是AU碱基对。

[0114] 在一个实施方案中，该正义链具有总计21个核苷酸，且该反义链具有总计23个核苷酸。

[0115] 在一个实施方案中，给药至表达包含六核苷酸重复的C9orf72靶标RNA的细胞之后，于24至48小时内，该dsRNA剂将包含六核苷酸重复的C9orf72靶标RNA的表达抑制至少10％、至少20％、至少30％、至少40％、至少50％或至少60％。

[0116] 在一个实施方案中，该dsRNA剂相对于成熟C9orf72信使RNA的表达而选择性地抑制包含六核苷酸重复的C9orf72靶标RNA的表达。

[0117] 在一个实施方案中，给药至表达成熟C9orf72信使RNA的细胞之后，于24至48小时内，该dsRNA剂将成熟C9orf72信使RNA的表达抑制少于50％、少于40％、少于30％、少于20％或少于10％。

[0118] 在一个实施方案中，给药至表达包含六核苷酸重复的C9orf72靶标RNA的细胞之后，于24至48小时内，该dsRNA剂减少二肽重复蛋白质合成。在一些实施方案中，给药至细胞之后，于24至48小时内，该dsRNA剂将二肽重复蛋白质合成减少至少10％、至少20％、至少30％、至少40％、至少50％、至少60％、至少70％或至少80％。

[0119] 本发明还提供细胞及药物组合物，其用于抑制编码C9orf72的基因的表达，其包含本发明的此类dsRNA剂。

[0120] 在一个实施方案中，该dsRNA剂处于非缓冲溶液诸如盐水或水中。

[0121] 在另一个实施方案中，该dsRNA剂处于缓冲溶液中，诸如包含醋酸盐、柠檬酸盐、醇溶谷蛋白、碳酸盐或磷酸盐或其任意组合的缓冲溶液中；或磷酸盐缓冲溶液(PBS)中。

[0122] 一方面，本发明提供抑制C9orf72(诸如含有六核苷酸重复的C9orf72 RNA，诸如包含六核苷酸重复的多个连续拷贝的C9orf72基因)于细胞例如神经元(诸如运动神经元)中的表达的方法，该方法包括使细胞与本发明的dsRNA剂或本发明的药物组合物接触，从而抑制C9orf72基因于细胞中的表达。

[0123] 另一方面，本发明提供减少细胞内的二肽重复蛋白质合成或二肽重复蛋白质集聚的方法。该方法包括将本发明的dsRNA剂引入细胞中，从而减少该细胞内的二肽重复蛋白质合成或二肽重复蛋白质集聚。

[0124] 另一方面，本发明提供减少细胞内的核和/或细胞质正义和/或反义C9orf72 RNA团簇的方法。该方法包括将本发明的dsRNA剂引入细胞中，从而减少该细胞内的核和/或细胞质正义和/或反义C9orf72RNA团簇。

[0125] 在一个实施方案中，细胞处于受试者内。

[0126] 在一个实施方案中，受试者是人。

[0127] 在一个实施方案中，受试者患有C9orf72相关疾患或处于发展出C9orf72相关疾患的风险中，诸如C9orf72-六核苷酸重复扩增相关疾病、病症或疾患。

[0128] 在一个实施方案中中，该C9orf72相关疾患选自由C9orf72肌萎缩侧索硬化/额颞叶痴呆或由于C9orf72扩增引起的亨廷顿样综合征所组成的组。

[0129] 在一个实施方案中，使该细胞与该dsRNA剂接触将C9orf72的表达抑制不超过50％。

[0130] 在一个实施方案中，给药至表达包含六核苷酸重复的C9orf72靶标RNA的细胞之后，于24至48小时内，该dsRNA剂将包含六核苷酸重复的C9orf72靶标mRNA的表达抑制至少10％、至少20％、至少30％、至少40％、至少50％或至少60％。

[0131] 在一些实施方案中，该dsRNA剂相对于成熟C9orf72信使RNA的表达而选择性地抑制包含六核苷酸重复的C9orf72靶标RNA的表达。在其他实施方案中，给药至表达成熟C9orf72信使RNA的细胞之后，于24至48小时内，该dsRNA剂将成熟C9orf72信使RNA的表达抑制少于50％、少于40％、少于30％、少于20％或少于10％。

[0132] 在一些实施方案中，该dsRNA剂减少该细胞内的二肽重复蛋白质合成或二肽重复蛋白质集聚。

[0133] 在一些实施方案中，该dsRNA剂减少细胞内的核和/或细胞质正义和/或反义C9orf72 RNA团簇。

[0134] 在一个实施方案中，抑制C9orf72的表达将该受试者血清中的C9orf72蛋白质水平降低不超过50％。

[0135] 在一些实施方案中，给药至细胞之后，于24至48小时内，该dsRNA剂将二肽重复蛋白质合成或二肽重复集聚减少至少10％、至少20％、至少30％、至少40％、至少50％、至少60％、至少70％或至少80％。

[0136] 一方面，本发明提供治疗患有将会受益于敲低靶标C9orf72 RNA的疾患(诸如C9orf72-六核苷酸重复扩增相关疾病、病症或疾患)的受试者的方法，包括向受试者给药治疗有效量的本发明的dsRNA剂或本发明的药物组合物，从而治疗患有将会受益于C9orf72表达减少的疾患的受试者。

[0137] 另一方面，本发明提供预防患有将会受益于C9orf72表达减少的疾患(诸如C9orf72-六核苷酸重复扩增相关疾病、病症或疾患)的受试者的至少一种症状的方法，包括向受试者给药预防有效量的本发明的dsRNA剂或本发明的药物组合物，从而预防患有将会受益于C9orf72表达减少的疾患的受试者的至少一种症状。

[0138] 在一个实施方案中，该疾患是C9orf72相关疾患。

[0139] 在一个实施方案中中，该C9orf72相关疾患选自由C9orf72肌萎缩侧索硬化/额颞叶痴呆或由于C9orf72扩增引起的亨廷顿样综合征所组成的组。

[0140] 在一个实施方案中，受试者是人。

[0141] 在一个实施方案中,将该剂给药至受试者造成C9orf72蛋白质累积的降低。

[0142] 在一些实施方案中,该方法减少受试者体内的二肽重复蛋白质合成或减少二肽重复蛋白质集聚。在一些实施方案中,该方法降低C9orf72靶标RNA于受试者体内的表达,该靶标RNA包含六核苷酸重复序列,该重复序列包含SEQ ID NO:1的多个连续拷贝。

[0143] 在一个实施方案中,dsRNA剂以约0.01mg/kg至约50mg/kg的剂量给药至该受试者。

[0144] 在一个实施方案中,dsRNA剂经皮下给药至受试者。

[0145] 在另一个实施方案中,daRNA剂经鞘内给药至受试者。

[0146] 在一个实施方案中,本发明的方法进一步包括测定来自该受试者的样本中C9orf72的水平。

[0147] 在一个实施方案中,该受试者样本中C9orf72的水平是血液、血清或脑脊液样本中C9orf72蛋白质水平。

[0148] 在一个实施方案中,本发明的方法进一步包括将附加治疗剂给药至受试者。

[0149] 一方面,本发明提供包含本发明的dsRNA剂或本发明的药物组合物的试剂盒(kit)。

[0150] 另一方面,本发明提供包含本发明的dsRNA剂或本发明的药物组合物的小瓶。

[0151] 又一方面,本发明提供包含本发明的dsRNA剂或本发明的药物组合物的注射器。

附图说明

[0152] 图1(不按比例)是野生型小鼠C9orf72基因座、野生型人类C9orf72基因座、以及人源化小鼠C9orf72基因座的示意图,其包含SEQ ID NO:1(GGGGCC)所示的六核苷酸序列(MAID8029a)的92次重复。小鼠序列由虚线以及带虚线的方框指示,而人类序列由实心黑色指示。外显子由方框表示。显示了六核苷酸重复序列的位置。

[0153] 图2(不按比例)是使用核酸酶剂将双链断裂引入邻近重复序列处以便使用重组机制来扩增位于靶标基因组基因座处的重复序列(诸如SEQ ID NO:1所示的位于C9orf72基因座处的六核苷酸序列)的示意图。

[0154] 图3(不按比例)是含C9orf72人源化、92x重复的ES细胞(MAID8029a)中定位在邻近该六核苷酸重复扩增序列处的八个导引RNA靶标序列的定位的示意图。

[0155] 图4示出于无细胞系统中,使用含有与人源化、含92x重复的等位基因(MAID8029a)相同序列的脂质体(8028Stvec)评估的八个导引RNA的裂解效率。

[0156] 图5(不按比例)是关于由将双链断裂引入该92x重复扩增序列的邻近5'端处(方框2)、邻近3'端处(方框3)、或紧邻5'端处及紧邻3'端处两处(方框1)而扩增该92x重复的示意图。黑色方框指示内源性小鼠序列;白色方框指示人源化区域。

[0157] 图6A(不按比例)是用来评估内源性C9orf72ES细胞克隆中的如SEQ ID NO:1中详述的六核苷酸序列的实例数目的传统双引物PCR的示意图。

[0158] 图6B(不按比例)是使用三种引物来评估内源性C9orf72ES细胞克隆中的如SEQ ID NO:1中详述的六核苷酸序列的实例数目的primePCR的示意图。

[0159] 图7A显示针对C9orf72基因座来评估在邻近该92x重复扩增序列的5'端处裂解后的重复区域尺寸的传统PCR的结果。经扩增的重复序列是由星号标记的。

[0160] 图7B显示针对C9orf72基因座来评估在邻近该92x重复扩增序列的3'端处裂解后

的重复区域尺寸的传统PCR的结果。经扩增的重复序列由星号标记的。

[0161] 图8显示primePCR的结果，以证实来自图7A的克隆中的一个以及来自图7B的克隆中的一个中重复序列的数目。使用亲本8029a(92x重复)克隆作为对照。图8显示毛细管电泳的结果。信号强度是于Y轴上，而PCR产物尺寸是于X轴上。读出是峰的数目。

[0162] 图9显示primePCR的结果，以证实来自图7A的克隆中的一个以及来自图7B的克隆中的一个中重复序列的数目。

[0163] 图10显示针对C9orf72基因座来评估在邻近该92x重复扩增序列的5'端处裂解后、在邻近该92x重复扩增序列的3'端处裂解后、或在邻近该92x重复扩增序列的5'端处及邻近3'端处两处裂解后的重复区域尺寸的传统PCR的结果。经扩增的重复序列是由箭头标记的。

[0164] 图11A(不按比例)显示于小鼠ES细胞中扩增92x重复的示意图：首先将双链断裂引入邻近该92x重复扩增序列的5'端处以生成第一经扩增克隆，随后，将双链断裂引入邻近该第一经扩增克隆的重复扩增序列的3'端处以生成第二经扩增克隆。人源化区域由标记及白色方框指示。指示5'DSB及3'DSB的定位，以及用于由PCR检测的引物的定位。

[0165] 图11B显示针对C9orf72基因组来评估250x重复扩增序列克隆在邻近250x重复扩增序列的3'处裂解后的第二扩增之后的重复区域尺寸的传统PCR的结果。经扩增重复序列及亲本重复序列是由带标记的箭头标记的。

[0166] 图12A至图12D是柱状图，显示在对于包含SEQ ID NO:1中详述的六核苷酸序列的3、92、250或300次重复的经修饰的C9orf72基因座为杂合者的胚胎干细胞来源运动神经元(ESMN)中，相对于包含SEQ ID NO:1中详述的六核苷酸序列的3次重复的ESMN中，来自C9orf72基因座的转录本的表达水平(y轴)(如由在每张图的顶部关于C9orf72的描述中所示的TAQMAN®定量逆转录偶联PCR(RT-qPCR)测定所测)，该转录本是外显子1A-外显子2剪接转录本(图12A)，该转录本是外显子1B-外显子2剪接转录本(图12B)，该转录本含有邻近外显子1A的内含子序列(图12C)，以及，该转录本保留邻近外显子1B的内含子序列(图12D)。

[0167] 图12E至图12H是柱状图，显示在对于包含SEQ ID NO:1中详述的六核苷酸序列的3、500或600次重复的经修饰的C9orf72基因座为杂合者的胚胎干细胞来源运动神经元(ESMN)中，相对于包含SEQ ID NO:1中详述的六核苷酸序列的3次重复的ESMN中，来自C9orf72基因座的转录本的表达水平(y轴)(如由在每张图的顶部关于C9orf72的描述中所示的TAQMAN®定量逆转录偶联PCR(RT-qPCR)测定所测)，该转录本是外显子1A-外显子2剪接转录本(图12E)，该转录本是外显子1B-外显子2剪接转录本(图12F)，该转录本含有邻近外显子1A的内含子序列(图12G)，以及，该转录本保留邻近外显子1B的内含子序列(图12H)。

[0168] 图13(上图)是来自对于经修饰的C9orf72基因座为杂合者的包含SEQ ID NO:1中详述的六核苷酸序列的3、92、300、500或600次重复的胚胎干细胞来源运动神经元(ESMN)的裂解液的westerm slot blot图像。包含0μg、1.25μg、2.5μg、5μg、10μg或20μg总蛋白质的裂解液使用抗poly-GlyPro抗体或抗poly-GlyAla抗体进行印迹。图13(下图)显示对该图在上部分的westerm slot blot的定量。

[0169] 图14A至图14B是柱状图，显示来自C9orf72的转录本的表达水平(y轴)(如由每张

图顶部的关于C9orf72的描述中所示的TAQMAN®定量逆转录偶联PCR(RT-qPCR)测定所测),该转录本含有对于包含SEQ ID NO:1中详述的六核苷酸序列的3、92、300、500或600次重复的经修饰的C9orf72基因座为杂合者的胚胎干细胞来源运动神经元(ESMN)中的邻近外显子1A的内含子序列,且表达水平是相对于包含SEQ ID NO:1中详述的六核苷酸序列的3次重复的ESMN(图14A);或者,该转录本是对于SEQ ID NO:1中详述的六核苷酸序列的3或300次重复的经修饰的(人源化)C9orf72基因座为杂合者的小鼠脑干及脊髓样本中的转录本,且表达水平是相对于包含SEQ ID NO:1中详述的六核苷酸序列的3次重复的小鼠脑干及脊髓样本(图14B)。

[0170] 图15A至图15D是柱状图,显示在胚胎干细胞来源运动神经元(ESMN)(其是类轴下(hypaxial-like)运动神经元或类四肢(limb-like)运动神经元,且对于包含SEQ ID NO:1中详述的六核苷酸序列的3、92、300、500或600次重复的经修饰的C9orf72基因座为杂合者)中,相对于包含SEQ ID NO:1中详述的六核苷酸序列的3次重复的ESMN中,来自C9orf72基因座的转录本的表达水平(y轴)(如由在每张图的顶部关于C9orf72的描述中所示的TAQMAN®定量逆转录偶联PCR(RT-qPCR)测定所测),该转录本是外显子1A-外显子2剪接转录本(图15A),该转录本是外显子1B-外显子2剪接转录本(图15B),该转录本含有邻近外显子1A的内含子序列(图15C),以及,该转录本保留邻近外显子1B的内含子序列(图15D)。

[0171] 图15E至图15F是柱状图,显示在胚胎干细胞来源运动神经元(ESMN)(其是类轴下运动神经元或类四肢运动神经元,且对于包含SEQ ID NO:1中详述的六核苷酸序列的3、92、300、500或600次重复的经修饰的C9orf72基因座为杂合者)中,相对于包含SEQ ID NO:1中详述的六核苷酸序列的3次重复的ESMN中,来自C9orf72基因座的转录本的表达水平(y轴)(如由在每张图的顶部关于C9orf72的描述中所示的TAQMAN®定量逆转录偶联PCR(RT-qPCR)测定所测),该转录本是未经剪接的前体转录本(图15E)或经剪接的C9orf72mRNA(图15F)。

[0172] 图16A至图16B是柱状图,显示在对于包含SEQ ID NO:1中详述的六核苷酸序列的300次重复的经修饰的C9orf72基因座为杂合者并且经靶向C9orf72的内含子1的siRNA、靶向经剪接的C9orf72mRNA的siRNA(外显子11(siRNA 1)、外显子5(siRNA 2及6)、外显子2(siRNA 3及5)、或外显子9(siRNA4))、或对照GAPDH siRNA处理24小时及48小时的胚胎干细胞来源运动神经元(ESMN)中,来自C9orf72基因座的转录本的表达水平(y轴)(如由在每张图的顶部关于C9orf72的描述中所示的TAQMAN®定量逆转录偶联PCR(RT-qPCR)测定所测),该转录本是外显子2-3剪接转录本(图16A),或者该转录本含有邻近外显子1A的内含子序列(图16B)。星号指示对应于图17A及图17Bwestern blot中的那些的样本。

[0173] 图17A是来自对于经修饰的C9orf72基因座为杂合者的包含SEQ ID NO:1中详述的六核苷酸序列的300次重复并且经靶向C9orf72的内含子1的siRNA、靶向经剪接的C9orf72mRNA的siRNA(外显子11(siRNA 1)、外显子5(siRNA 2)或外显子2(siRNA 3))或对照GAPDH siRNA处理48小时后的胚胎干细胞来源运动神经元(ESMN)的裂解液的western slot blot图像。包含1.25μg或2.5μg总蛋白质的裂解液使用抗poly-GlyAla抗体进行印迹。图17B显示对图17A中的western slot blot的定量。

[0174] 图18A是靶向C9orf72的内含子1A的siRNA的示意图(不按比例)。

[0175] 图18B是跨越C9orf72 pre-mRNA的不同节段的TAQMAN®定量PCR测定A、B、C、D、E、F、G、H、I及J示意图(不按比例)。引物及探针的序列显示于表8中。

[0176] 图19A至图19C是柱状图,显示使用靶向外显子1A的siRNA(siRNA E1A#1、E1A#2及E1A#3)、靶向位于六核苷酸重复上游的内含子1A的siRNA(siRNA I1A#2、I1A#4、I1A#6及I1A#10)或靶向位于六核苷酸重复下游及外显子1B下游的内含子1B的siRNA(siRNA I1B#1及I1B#6)在包含经修饰的C9orf72基因座(其包含SEQ ID NO:1中详述的六核苷酸序列的300次重复)的小鼠胚胎干细胞中转染24小时后,来自C9orf72基因座的转录本的相对表达水平(如由TAQMAN®定量逆转录偶联PCR(RT-qPCR)测定而测定者)(y轴),该转录本含有C9orf72内含子1A序列(图19A;测定F),该转录本是C9orf72外显子2-3剪接转录本(图19B),或者该转录本是C9orf72外显子1a-2剪接转录本(图19C;测定B)。使用Gapdh siRNA作为对照,并且使用模拟作为阴性对照。

[0177] 图20A至图20C是柱状图,显示使用靶向外显子1A的siRNA(siRNA E1A#1、E1A#2及E1A#3)、靶向位于六核苷酸重复上游的内含子1A的siRNA(siRNA I1A#2、I1A#4、I1A#6及I1A#10)或靶向位于六核苷酸重复下游及外显子1B下游的内含子1B的siRNA(siRNA I1B#1及I1B#6)在包含经修饰的C9orf72基因座(其包含SEQ ID NO:1中详述的六核苷酸序列的300次重复)的小鼠胚胎干细胞中转染48小时后,来自C9orf72基因座的转录本的相对表达水平(如由TAQMAN®定量逆转录偶联PCR(RT-qPCR)测定而测定者)(y轴),该转录本含有C9orf72内含子1A(图20A;测定F),该转录本是C9orf72外显子2-3剪接转录本(图20B),或者该转录本是C9orf72外显子1a-2剪接转录本(图20C;测定B)。使用Gapdh siRNA作为对照,并且使用模拟作为阴性对照。

[0178] 图21A至图21C是柱状图,显示使用靶向外显子1A的siRNA(siRNA E1A#1、E1A#2及E1A#3)、靶向位于六核苷酸重复上游的内含子1A的siRNA(siRNA I1A#2、I1A#4、I1A#6及I1A#10)或靶向位于六核苷酸重复下游及外显子1B下游的内含子1B的siRNA(siRNA I1B#1及I1B#6)在包含经修饰的C9orf72基因座(其包含SEQ ID NO:1中详述的六核苷酸序列的300次重复)的小鼠胚胎干细胞中转染72小时后,来自C9orf72基因座的转录本的相对表达水平(如由TAQMAN®定量逆转录偶联PCR(RT-qPCR)测定而测定者)(y轴),该转录本含有C9orf72内含子1A(图21A;测定F),该转录本是C9orf72外显子2-3剪接转录本(图21B),或者该转录本是C9orf72外显子1a-2剪接转录本(图21C;测定B)。使用Gapdh siRNA作为对照,并且使用模拟作为阴性对照。

[0179] 图22A至图22E是柱状图,显示使用靶向成熟C9orf72 mRNA的siRNA(siRNA#1)或靶向位于六核苷酸重复上游的内含子1A的siRNA(siRNA#12)在包含经修饰的C9orf72基因座(其包含SEQ ID NO:1中详述的六核苷酸序列的300次重复)的小鼠胚胎干细胞中电穿孔24小时后,来自C9orf72基因座的转录本的相对表达水平(如由TAQMAN®定量逆转录偶联PCR(RT-qPCR)测定而测定者)(y轴),该转录本含有C9orf72内含子1A序列(图22A;测定F),该转录本含有连接至外显子1A的C9orf72内含子1A序列(图22B;测定E),该转录本是C9orf72外显子1a-2剪接的转录本(图22C;测定B),该转录本是C9orf72外显子1b-2剪接的转录本(图22D;测定C),或者该转录本是C9orf72外显子5-6剪接的转录本(图22E;测定D)。使用Gapdh siRNA及模拟样本作为阴性对照。

[0180] 图23A至图23E是柱状图，显示使用靶向成熟C9orf72 mRNA的siRNA(siRNA#1)或靶向位于六核苷酸重复上游的内含子1A的siRNA(siRNA#12)在包含经修饰的C9orf72基因座(其包含SEQ ID NO:1中详述的六核苷酸序列的300次重复)的小鼠胚胎干细胞中电穿孔48小时后，来自C9orf72基因座的转录本的相对表达水平(如由TAQMAN®定量逆转录偶联PCR(RT-qPCR)测定而测定者)(y轴)，该转录本含有C9orf72内含子1A序列(图23A;测定F)，该转录本含有连接至外显子1A的C9orf72内含子1A序列(图23B;测定E)，该转录本是C9orf72外显子1a-2剪接的转录本(图23C;测定B)，该转录本是C9orf72外显子1b-2剪接的转录本(图23D;测定C)，或者该转录本是C9orf72外显子5-6剪接的转录本(图23E;测定D)。使用Gapdh siRNA及模拟样本作为阴性对照。

具体实施方式

[0181] 本公开提供RNAi组合物，其影响C9orf72基因(诸如具有经扩增的GGGGCC(G_4C_2)重复的C9orf72基因)的RNA转录本的RNA诱导沉默复合物(RISC)介导的裂解。C9orf72基因可以处于细胞内，例如，受试者诸如人体内的细胞内。这些iRNA的使用使得哺乳动物体内相应基因(C9orf72基因)的mRNA的靶向降解成为可能。

[0182] 本发明的iRNA已经设计为靶向C9orf72靶标RNA，例如，于该基因的内含子中具有经扩增的GGGGCC六核苷酸重复并且具有核苷酸修饰的组合的C9orf72靶标RNA。这些剂可靶向成熟C9orf72mRNA(内含子被剔除的mRNA)或C9orf7 RNA前体(含有内含子的RNA)。本发明的iRNA可将C9orf7基因(例如成熟mRNA)的表达抑制不超过约50％，并且降低含正义及反义的团簇及异常二肽重复(DPR)蛋白质(poly(GA)、poly(GR)、poly(GP)、poly(PA)及poly(PR))的水平，和/或将C9orf7基因(例如，前体mRNA)的表达抑制超过约50％，并且降低含正义及反义的团簇及异常二肽重复(DPR)蛋白质(poly(GA)、poly(GR)、poly(GP)、poly(PA)及poly(PR))的水平。不试图受限于理论，据信前述特性与这些iRNA中的特定靶标位点或特定修饰的组合或亚组合赋予本发明的iRNA以改善的功效、稳定性、效力、耐久性及安全性。

[0183] 因此，本公开还提供使用本公开的RNAi组合物用于抑制C9orf72基因表达或用于治疗患有将会受益于C9orf72基因表达抑制或降低的疾患的受试者的方法，例如C9orf72相关疾病，该疾病为例如与C9orf72基因的内含子中经扩增的GGGGCC六核苷酸重复相关的疾病，诸如C9orf72肌萎缩侧索硬化/额颞叶痴呆，或亨廷顿氏病，例如，由于C9orf72扩增引起的亨廷顿样综合征。

[0184] 本公开的RNAi剂包括RNA链(反义链)，该链是具有约30个核苷酸或更短的长度的区域，例如15至30、15至29、15至28、15至27、15至26、15至25、15至24、15至23、15至22、15至21、15至20、15至19、15至18、15至17、18至30、18至29、18至28、18至27、18至26、18至25、18至24、18至23、18至22、18至21、18至20、19至30、19至29、19至28、19至27、19至26、19至25、19至24、19至23、19至22、19至21、19至20、20至30、20至29、20至28、20至27、20至26、20至25、20至24、20至23、20至22、20至21、21至30、21至29、21至28、21至27、21至26、21至25、21至24、21至23或21至22个核苷酸的长度，该区域实质上与C9orf72基因的靶标RNA转录本的至少一部分例如C9orf72内含子互补。在某些实施方案中，本公开的RNAi剂包括具有约21至23个核苷酸的长度的区域的RNA链(反义链)，该区域与C9orf72基因的靶标RNA转录本的至少一部分例如C9orf72内含子基本上互补。

[0185] 在某些实施方案中，本公开的RNAi剂包括RNA链(反义链)，该链可包括更长的长度，例如多达66个核苷酸，例如36至66、26至36、25至36、31至60、22至43、27至53个核苷酸的长度，具有一个实质上与C9orf72基因的mRNA转录本的至少一部分互补的至少19个连续核苷酸的区域。这些具有长度较长的反义链的RNAi剂优选包括长度为20至60个核苷酸的第二RNA链(正义链)，其中该正义链与该反义链是形成18至30个连续核苷酸的双链体。

[0186] 这些RNAi剂的使用使得对哺乳动物体内的C9orf72基因的靶标RNA的靶向降解成为可能。因此，包括这些RNAi剂的方法及组合物可用于治疗将会受益于靶标C9orf72 RNA的敲低、正常C9orf72蛋白质的降低和/或自致病性六核苷酸重复扩增生成的致病性二肽重复蛋白质的降低的受试者，诸如患有C9orf72相关疾病诸如C9orf72肌萎缩侧索硬化/额颞叶痴呆或亨廷顿氏病例如由于C9orf72扩增引起的亨廷顿样综合征的受试者。

[0187] 下述的具体实施方式披露了如何制作及使用含有用以抑制C9orf72基因表达的RNAi剂的组合物，以及治疗患有将会从抑制或降低这些基因的表达中受益的疾病及疾患的受试者的方法。

[0188] I.定义

[0189] 为了更容易地理解本公开，首先定义某些术语。此外，应注意，无论何时，当记载参数的数值或数值范围，这些数值及处于这些所引用的数值中间的范围也作为本公开的一部分。

[0190] 本文中使用的冠词“一个/种”是指该冠词的语法对象的一个/种或超过一个/种(即，至少一个/种)。例如，“一个组件”意指一个组件或超过一个组件如多个组件。

[0191] 本文中使用的术语“包括”意指“包括但不限于”，且与后者可互换地使用。除非上下文明确地另有所指，否则本文中使用的术语“或”意指“和/或”，且与后者可互换地使用。

[0192] 本文中使用的术语“约”意指处于本领域中的典型公差范围内。例如，“约”可理解为与均值偏离2个标准偏差。在某些实施方案中，“约”意指±10％。在某些实施方案中，“约”意指±5％。当“约”存在于一系列数字或范围之前时，理解为“约”可修饰该一系列数字或范围中的每一个。

[0193] 处于数字或一系列数字之前的术语“至少”理解为包括与该术语“至少”相邻的数字，以及后面的全部数字或逻辑上可包括的整数，如从语境中明显可知的。例如，核酸分子中的核苷酸的数目必需为整数。例如，“21个核苷酸的核酸分子的至少18个核苷酸”意指18、19、20或21个核苷酸具有所指示的特性。当“至少”存在于一系列数字或范围之前时，理解为“至少”可修饰该一系列数字或范围中的每一个。

[0194] 如本文所用，“不超过”或“少于”理解为与短语相邻的数值以及逻辑上更低的数值或整数，如从语境中逻辑上可推知的，到零为止。例如，具有“不超过2个核苷酸”的突出端的双链体具有2、1或0个核苷酸的突出端。当“不超过”存在于一系列数字或范围之前时，理解为“不超过”可修饰该一系列数字或范围中的每一个。

[0195] 如本文所用，检测方法可包括确定所存在的分析物的量低于该方法的检测水平。

[0196] 在所指示的靶标位点与正义或反义链的核苷酸序列之间存在矛盾的情况下，以所指示的序列为准。

[0197] 在化学结构与化学名称之间存在矛盾的情况下，以化学结构为准。

[0198] “包含”或“包括”一种或多种所记载组件的组合物或方法可包括未具体记载的其

他组件。例如，“包含”或“包括”蛋白质的组合物可含有单独的蛋白质或与其他成分组合的蛋白质。连接短语“基本上由...组成”意指权利要求的范围解释为涵盖该权利要求中记载的指定组件以及在材料上不影响请求保护的发明的基本特征及新颖特征的那些。因此，当用于本发明的权利要求中时，术语“基本上由...组成”不试图解释为等同于“包含”。

[0199] “任选的”或“任选地”意指随后描述的事件或环境可出现或不出现，并且该描述包括该事件或环境出现的实例以及该事件或环境不出现的实例。

[0200] 术语“C9orf72”基因，也称为“C9orf72-SMCR8复合物亚单元”、“鸟嘌呤核苷酸交换C9orf72”、“染色体9开放阅读框72”、“蛋白质C9orf72”、“DENNL72”、“FTDALS1”、“ALSFTD”及“FTDALS”，是指编码内体运输调控中涉及的熟知蛋白质即，C9orf72的基因。已经证明，C9orf72蛋白质与自噬及胞吞转运中涉及的Rab蛋白质交互作用。GGGGCC重复在位于这一基因转录本的交替5’外显子之间的内含子性序列中从约2至约22个拷贝扩增为约700至约1600个拷贝相关于C9orf72肌萎缩侧索硬化/额颞叶痴呆及亨廷顿氏病，例如，由于C9orf72扩增引起的亨廷顿样综合征。替代剪接导致编码不同同种型的多个转录本变体。

[0201] C9orf72的示例性核苷酸和氨基酸序列可见于，例如，GenBank登录号NM_001256054.2(智人C9orf72，SEQ ID NO:121，反向互补序列SEQ ID NO:125)；GenBank登录号XM_005581570.2(食蟹猴C9orf72，SEQ ID NO:122，反向互补序列SEQ ID NO:126)；GenBank登录号NM_001081343.2(小鼠C9orf72，SEQ ID NO:123，反向互补序列SEQ ID NO:127)；以及GenBank登录号NM_001007702.1(大鼠C9orf72，SEQ ID NO:124，反向互补序列SEQ ID NO:128)。

[0202] 人类C9orf72的另外核苷酸及氨基酸序列可见于，例如，GenBank登录号NM_145005.6，转录本变体1(SEQ ID NO:129，反向互补序列SEQ ID NO:130)以及NM_018325.5，转录本变体2(SEQ ID NO:131，反向互补序列SEQ ID NO:132)。

[0203] 带有C9orf72基因的人类染色体9基因组区域的核苷酸序列可见于，例如，可于GenBank获得的人类基因组参考序列联盟第38版(Genome Reference Consortium Human Build 38，也称为人类基因组第38版或GRCh38)。带有C9orf72基因的人类染色体9基因组区域的核苷酸序列也可见于，例如，GenBank登录号NC_000009.12(SEQ ID NO:133，反向互补序列SEQ ID NO:134)，对应于人类染色体9的核苷酸27546545至27573866。人c9orf72基因的核苷酸序列可见于，例如，GenBank登录号NG_031977.1。

[0204] C9orf72序列的进一步示例可见于可公开获得的数据库，例如，GenBank、OMIM及UniProt。

[0205] C9orf72的另外信息可见于，例如，www.ncbi.nlm.nih.gov/gene/203228。如本文中所用，术语C9orf72还指C9orf72基因的包括临床变体数据库中提供的变体的版本，例如，于www.ncbi.nlm.nih.gov/clinvar/?term=NM_001256054.2。

[0206] 自递交本申请之日起，前述GenBank登录号及Gene数据库号的每一者的全部内容通过援引并入本文。

[0207] 如本文中所用，“靶标序列”是指在C9orf72基因的转录过程中形成的mRNA分子的核苷酸序列的连续部分，包括作为初级转录产物的RNA加工产物的mRNA。在一个实施方案中，该序列的靶标部分将至少长至足以用作RNAi引导的裂解的底物，该裂解位于在C9orf72基因的转录过程中形成的mRNA分子的核苷酸序列的部分处或邻近该处。

[0208] C9orf72 mRNA(靶标C9orf72 RNA)是自C9orf72基因转录的RNA。C9orf72 mRNA包括C9orf72成熟mRNA、C9orf72前体mRNA、或其任何部分(例如,剔除内含子区域或替代性经剪接的RNA)。C9orf72成熟mRNA是其内含子已经移除(剔除)的C9orf72 mRNA,并且C9orf72蛋白质是自其翻译。C9orf72前体mRNA是至少一个内含子尤其是第一内含子(内含子1)尚未被移除的C9orf72 mRNA。

[0209] C9orf72蛋白质包括自C9orf72 mRNA表达的任何蛋白质。C9orf72蛋白质包括自C9orf72成熟mRNA表达的蛋白质,以及自含有六核苷酸重复的C9orf72 RNA进行重复序列相关的非AUG(AUG)翻译所导致的二肽重复蛋白质(例如,聚(甘氨酸-丙氨酸)、聚(甘氨酸-脯氨酸)、聚(甘氨酸-精氨酸)、聚(丙氨酸-脯氨酸)及聚(脯氨酸-精氨酸))。

[0210] C9orf72靶标RNA可包括包含SEQ ID NO:1的多个连续拷贝的六核苷酸重复。C9orf72靶标RNA可以是,例如,具有致病性六核苷酸重复扩增(具有,例如,至少约30个、至少约35个、至少约40个、至少约50个、至少约60个、至少约70个、至少约80个、至少约100个、至少约200个、至少约300个、至少约400、或至少约500个六核苷酸重复的拷贝)的那种。

[0211] 靶标序列可以是约15至30个核苷酸的长度。例如,靶标序列可以是约15至30个核苷酸的长度,如15至29、15至28、15至27、15至26、15至25、15至24、15至23、15至22、15至21、15至20、15至19、15至18、15至17、18至30、18至29、18至28、18至27、18至26、18至25、18至24、18至23、18至22、18至21、18至20、19至30、19至29、19至28、19至27、19至26、19至25、19至24、19至23、19至22、19至21、19至20、20至30、20至29、20至28、20至27、20至26、20至25、20至24、20至23、20至22、20至21、21至30、21至29、21至28、21至27、21至26、21至25、21至24、21至23、或21至22个核苷酸的长度。在某些实施方案中,靶标序列是19至23个核苷酸的长度,任选地是21至23个核苷酸的长度。上文引用的范围及长度之间的范围及长度也视为本公开的一部分。

[0212] 如本文中所使用,术语“包含序列的链”是指包含核苷酸的链的寡核苷酸,该核苷酸由使用标准核苷酸命名法指代的序列而描述。

[0213] 在经修饰或未修饰的核苷酸的语境中,“G”、“C”、“A”、“T”及“U”各自通常分别表示含有鸟嘌呤、胞嘧啶、腺嘌呤、胸腺嘧啶及尿嘧啶作为碱基的核苷酸。然而,应理解,术语“核糖核苷酸”或“核苷酸”还可指代经修饰的核苷酸,如下文进一步描述的,或替代替换部分(参见,例如,表11)。本领域技术人员熟知,鸟嘌呤、胞嘧啶、腺嘌呤、胸腺嘧啶及尿嘧啶可经由其他部分置换而基本上不改变包含携带此置换部分的核苷酸的寡核苷酸的碱基配对特性。例如而非限制性的,包含肌苷作为其碱基的核苷酸可与含有腺嘌呤、胞嘧啶或鸟嘌呤的核苷酸进行碱基配对。因此,在本公开提出的dsRNA的核苷酸序列中,含有尿嘧啶、鸟嘌呤或腺嘌呤的核苷酸可置换为含有例如肌苷的核苷酸。在另一示例中,寡核苷酸中任何位置的腺嘌呤及胞嘧啶可分别替换为鸟嘌呤及尿嘧啶,以与靶标mRNA形成G-U Wobble碱基配对。含有此类置换部分的序列适用于本公开提出的组合物及方法。

[0214] 如本文中可互换使用,术语“iRNA”、“RNAi剂”、“iRNA剂”、“RNA干扰剂”是指含有如本文中定义的术语的RNA,且其经由RNA诱导沉默复合物(RISC)途径而介导RNA转录本的靶向裂解。RNA干扰(RNAi)是引导mRNA的序列特异性降解的过程。RNAi在细胞例如受试者诸如哺乳动物受试者体内的细胞内调控(例如抑制)C9orf72、C9orf72-相关转录本或C9orf72相关肽的表达。

[0215] 在一个实施方案中，本公开的RNAi剂包括单链RNAi，其与靶标RNA序列例如C9orf72靶标mRNA相互作用，以引导该靶标RNA的裂解。不希望受理论束缚，据信被引入细胞内的长双链RNA借由被称为切丁酶(Dicer)的III型核酸内切酶而破碎为包含正义链及反义链的双链短干扰RNA(siRNA)(Sharp et al.(2001)Genes Dev.15:485)。切丁酶，类核酸酶III酶，将这些dsRNA加工为19至23个碱基对的短干扰RNA，该短干扰RNA的特征是具有两个碱基的3'突出端(Bernstein,et al.,(2001)Nature 409:363)。随后，这些siRNA被并入RNA诱导沉默复合物(RISC)内，在该处，一种或多种解旋酶使该siRNA双链体解卷曲，使得互补反义链能够引导靶标识别(Nykanen,et al.,(2001)Cell 107:309)。当结合至适宜的靶标mRNA时，RISC内的一种或多种核酸内切酶裂解靶标以诱导沉默(Elbashir,et al.,(2001)Genes Dev.15:188)。因此，一方面，本公开涉及单链RNA(ssRNA)(siRNA双链体的反义链)，其在细胞内生成且促进RISC复合物的形成以有效沉默靶标基因即C9orf72基因。据此，本文中，术语"siRNA"还用以是指上文所述的RNAi。

[0216] 在另一实施方案中，该RNAi剂可以是单链RNA，其被引入细胞或有机体内以抑制靶标mRNA。单链RNAi剂结合至RISC核酸内切酶Argonaute 2，其随后裂解靶标mRNA。该单链siRNA通常是15至30个核苷酸且经化学修饰。单链RNA的设计及测试描述于美国专利第8,101,348号及Lima et al.,(2012)Cell 150:883-894中，其各自的全部内容通过援引而并入本文。本文中描述的任何反义核苷酸序列可用作本文中描述的单链siRNA或通过Lima et al.,(2012)Cell150:883-894中描述的方法进行化学修饰。

[0217] 在另一实施方案中，用于本公开的组合物及方法中的"RNAi剂"是双链RNA，且在本文中称为"双链RNAi剂"、"双链RNA(dsRNA)分子"、"dsRNA分子"或"dsRNA"。术语"dsRNA"是指核糖核酸分子的复合物，其具有包含两个反平行且基本上互补的核酸链的双链体结构，这两个核酸链称为具有相对于靶标RNA即C9orf72基因的"正义"取向及"反义"取向。在本公开的一些实施方案中，双链RNA(dsRNA)通过转录后基因沉默机制而触发靶标RNA例如mRNA的降解，本文中，该机制称为RNA干扰或RNAi。

[0218] 本文所描述的dsRNA剂可不同于(即，不包括)反义寡核苷酸(ASO)或间隙序列(gapmer)反义寡核苷酸(ASO)。

[0219] 通常，dsRNA分子可包括核糖核苷酸，但如本文中所详述，每条链或两条链还可包括一个或多个非核糖核苷酸，例如脱氧核糖核苷酸、经修饰的核苷酸。此外，如本说明书中所用，"RNAi剂"可包括具有化学修饰的核糖核苷酸;RNAi剂可包括位于多个核苷酸处的实质性修饰。如本文中所用，术语"经修饰的核苷酸"是指独立具有经修饰的糖部分、经修饰的核苷酸间键、或经修饰的核碱基的核苷酸。因此，术语"经修饰的核苷酸"涵盖例如官能基团或原子至核苷酸间键、糖部分或核碱基的置换、添加或移除。适用于本公开的剂中的修饰包括本文中所披露或本领域中已知的全部类型的修饰。出于本说明书及权利要求书的目的，任何此类修饰，如在siRNA类型分子中所用的那些，为"RNAi剂"所涵盖。

[0220] 在本公开的某些实施方案中，将脱氧核苷酸(若存在)包括在RNAi剂中可被视为构建经修饰的核苷酸。

[0221] 该双链体区域可以是允许所希望的靶标RNA经由RISC途径而特异性降解的任何长度，且长度范围可以是约15至36个碱基对，例如，约15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35或36个碱基对的长度，诸如约15至30、15至29、15至28、15

至27、15至26、15至25、15至24、15至23、15至22、15至21、15至20、15至19、15至18、15至17、18至30、18至29、18至28、18至27、18至26、18至25、18至24、18至23、18至22、18至21、18至20、19至30、19至29、19至28、19至27、19至26、19至25、19至24、19至23、19至22、19至21、19至20、20至30、20至29、20至28、20至27、20至26、20至25、20至24,20至23、20至22、20至21、21至30、21至29、21至28、21至27、21至26、21至25、21至24、21至23、或21至22个碱基对的长度。在某些实施方案中,双链体区域是19至21个碱基对的长度,例如,21个碱基对的长度。上文列举的范围及长度之间的范围及长度也视为本公开的一部分。

[0222] 形成该双链体结构的两条链可以是一个较大RNA分子的不同部分,或它们可以是独立的RNA分子。若这两条链是一个较大分子的部分,且因此通过一条链的3'末端与形成该双链体结构的相对另一条链的5'末端之间的未中断核苷酸链而连接,则该连接RNA链称为"发夹环"。发夹环可包含至少一个未配对的核苷酸。在一些实施方案中,该发夹环可包含至少4、至少5、至少6、至少7、至少8、至少9、至少10、至少20、至少23或更多个未配对的核苷酸或未被引导至dsRNA靶标位点的核苷酸。在一些实施方案中,发夹环可以是10个或更少核苷酸。在一些实施方案中,发夹环可以是8个或更少未配对的核苷酸。在一些实施方案中,发夹环可以是4至10个未配对的核苷酸。在一些实施方案中,发夹环可以是4至8个核苷酸。

[0223] 若dsRNA的两条基本上互补的链由独立的RNA分子构成,则那些分子不必但可以共价连接。在某些实施方案中,若这两条链是通过除在一条链的3'端与形成该双链体结构的相对另一条链的5'端之间的未中断核苷酸链以外的手段共价连接,则该连接结构称为"接头"(但应注意,本文中他处定义的某些其他结构也可称为"接头")。RNA链可具有相同或不同数目的核苷酸。碱基对的最大数目是dsRNA的最短链中的核苷酸数减去双链体中存在的任何突出端。RNAi除了包含双链体结构外,还可包含一个或多个核苷酸突出端。在RNAi剂的一个实施方案中,至少一条链包含至少1个核苷酸的3'突出端。在另一实施方案中,至少一条链包含具有至少2个核苷酸,例如2、3、4、5、6、7、9、10、11、12、13、14或15个核苷酸的3'突出端。在其他实施方案中,RNAi剂的至少一条链包含具有至少1个核苷酸的5'突出端。在某些实施方案中,至少一条链包含具有至少2个核苷酸,例如2、5、4、5、6、7、9、10、11、12、13、14或15个核苷酸的5'突出端。在又其他实施方案中,RNAi剂的一条链的3'及5'端包含具有至少1个核苷酸的突出端。

[0224] 在一个实施方案中,本公开的RNAi剂是dsRNA,其每一条链独立地包含19至23个核苷酸,其与靶标RNA序列如C9orf72靶标mRNA序列相互作用以引导靶标RNA的裂解。

[0225] 在一些实施方案中,本发明的iRNA剂是24至30个核苷酸的dsRNA,其与靶标RNA序列例如C9orf72靶标mRNA序列相互作用以引导靶标RNA的裂解。

[0226] 如本文中所用,术语"核苷酸突出端"是指从RNAi剂例如dsRNA的双链体结构凸出的至少一个未配对的核苷酸。例如,当dsRNA的一条链的3'端延伸超过另一条链的5'端(反之亦然),则存在核苷酸突出端。dsRNA可包含具有至少一个核苷酸的突出端;或者突出端可包含至少两个核苷酸、至少三个核苷酸、至少四个核苷酸、至少五个核苷酸或更多个。核苷酸突出端可包含核苷酸/核苷类似物或由其组成,其中该核苷酸/核苷类似物包括脱氧核苷酸/核苷。突出端可位于正义链、反义链或其任何组合上。此外,突出端的核苷酸可存在于dsRNA的反义链或正义链的5'端、3'端或两端。

[0227] 在一个实施方案中,dsRNA的反义链具有突出端在3'端或5'端的1至10个核苷酸,

例如,1、2、3、4、5、6、7、8、9、或10个核苷酸。在一个实施方案中,dsRNA的正义链具有突出端在3’端或5’端的1至10个核苷酸,例如,1、2、3、4、5、6、7、8、9、或10个核苷酸。在另一实施方案中,该突出端中的一个或多个核苷酸被替换为核苷硫代磷酸酯。

[0228] 在某些实施方案中,该dsRNA的反义链具有突出端在3’端或5’端的1至10个核苷酸,例如0至3、1至3、2至4、2至5、4至10、5至10个,例如1、2、3、4、5、6、7、8、9、或10个核苷酸。在一个实施方案中,dsRNA的正义链具有突出端在3’端或5’端的1至10个核苷酸,例如,1、2、3、4、5、6、7、8、9、或10个核苷酸。在另一实施方案中,该突出端中的一个或多个核苷酸被替换为核苷硫代磷酸酯。

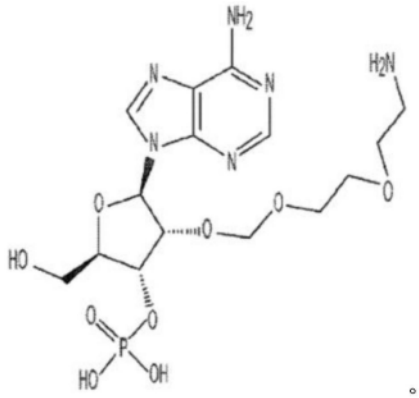
[0229] 在某些实施方案中,位于正义链或反义链的突出端可包括长于10个核苷酸的延伸长度,例如,1至30个核苷酸、2至30个核苷酸、10至30个核苷酸或10至15个核苷酸的长度。在某些实施方案中,延伸的突出端位于双链体的正义链。在某些实施方案中,延伸的突出端存在于双链体的正义链的3’端。在某些实施方案中,延伸的突出端存在于双链体的正义链的5’端。在某些实施方案中,延伸的突出端位于双链体的反义链。在某些实施方案中,延伸的突出端存在于双链体的反义链的3’端。在某些实施方案中,延伸的突出端存在于双链体的反义链的5’端。在某些实施方案中,突出端中的一个或多个核苷酸被替换为核苷硫代磷酸酯。在某些实施方案中,突出端包括自身互补的部分,使得该突出端能够形成在生理学条件下稳定的发夹结构。

[0230] 在某些实施方案中,至少一条链的至少一端延伸超出双链体靶向区域,包括其中这些链中的一个包括热力学稳定化四环结构的结构(参见,例如,美国专利第8,513,207号及第8,927,705号以及WO2010033225,其各自的全部内容通过援引并入本文)。此类结构可包括单链延伸(于分子的一侧或两侧)以及双链延伸。

[0231] 在某些实施方案中,正义链的3’端及反义链的5’端通过包含核糖核苷酸、脱氧核糖核苷酸或两者的多核苷酸序列连接,任选地,其中该多核苷酸序列包含四环序列。在某些实施方案中,正义链是25至35个核苷酸的长度。

[0232] 四环可含有核糖核苷酸、脱氧核糖核苷酸、经修饰的核苷酸及其组合。典型地,四环具有4至5个核苷酸。在一些实施方案中,环包含如GAAA所述的序列。在一些实施方案中,环(GAAA)的核苷酸中的至少一个包含核苷酸修饰。在一些实施方案中,经修饰的核苷酸包含2’-修饰。在一些实施方案中,2’-修饰是选自由2’-氨基乙基、2’-氟、2’-O-甲基、2’-O-甲氧基乙基、2’-氨基二乙氧基甲醇、2’-adem及2’-脱氧-2’氟--d-阿拉伯糖核酸所组成的组的修饰。在一些实施方案中,环的全部核苷酸皆经修饰。在一些实施方案中,GAAA序列中的G包含2’-OH。在一些实施方案中,GAAA序列中的每个核苷酸皆包含2’-O-甲基修饰。在一些实施方案中,GAAA序列中的每个A皆包含2’-OH,并且GAAA序列中的G包含2’-O-甲基修饰。在优选的实施方案中,在一些实施方案中,GAAA序列中的每个A皆包含2’-O-甲氧基乙基(MOE)修饰,并且GAAA序列中的G包含2’-O-甲基修饰;或GAAA序列中的每个A包含2’-adem修饰,并且GAAA序列中的G包含2’-O-甲基修饰。参见,例如,PCT公布WO 2020/206350,其全部内容通过援引而并入本文。

[0233] 示例性的经2’-adem修饰的核苷酸如下所示:

[0234] 

。

[0235] 如本文中所用,关于dsRNA的术语“钝”或“钝端”意指,在dsRNA的给定端无未配对的核苷酸或核苷酸类似物,即,无核苷酸突出端。dsRNA的一端或两端可以是钝端。若dsRNA的两端都是钝端,该dsRNA称为钝端的。清楚起见,“钝端的”dsRNA是两端都是钝的dsRNA,即,在分子的任一端都无核苷酸突出端。最常见的此类分子将在整个长度上是双链的。

[0236] 术语“反义链”或“导引链”是指RNAi剂例如dsRNA的链,其包括与靶标序列如C9orf72 mRNA基本上互补的区域。

[0237] 如本文中所用,术语“互补的区域”是指反义链上的与本文中定义的序列(例如靶标序列例如C9orf72核苷酸序列)基本上互补的区域。若互补的区域与靶标序列不完全互补,则错配可存在于分子的中间区域或末端区域。通常,最能被容忍的错配存在于末端区域内,例如,RNAi剂的5’末端或3’末端的5、4、3或2个核苷酸内。在一些实施方案中,本发明的双链RNA剂包括位于反义链中的核苷酸错配。在一些实施方案中,本发明的双链RNA剂的反义链包括不超过4个与靶标mRNA的错配,例如,反义链包括4、3、2、1或0个与靶标mRNA的错配。在一些实施方案中,本发明的双链RNA剂的反义链包括不超过4个与正义链的错配,例如,反义链包括4、3、2、1或0个与正义链的错配。在一些实施方案中,本发明的双链RNA剂包括位于正义链的核苷酸错配。在一些实施方案中,本发明的双链RNA剂的正义链包括不超过4个与反义链的错配,例如,正义链包括4、3、2、1或0个与反义链的错配。在一些实施方案中,核苷酸错配位于例如自iRNA的3’端的5、4、3个核苷酸内。在另一实施方案中,核苷酸错配是例如位于iRNA剂的3’-末端核苷酸中。在一些实施方案中,错配不处于种子区域中。

[0238] 因此,本文中所描述的RNAi剂可含有一个或多个与靶标序列的错配。在一个实施方案中,本文中所描述的RNAi剂含有不超过3个错配(即,3、2、1或0个错配)。在一个实施方案中,本文中所描述的RNAi剂含有不超过2个错配。在一个实施方案中,本文中所描述的RNAi剂含有不超过1个错配。在一个实施方案中,本文中所描述的RNAi剂含有0个错配。在某些实施方案中,如果RNAi剂的反义链含有与靶标序列的错配,则该错配优选可被限定在从互补区域的5’端或3’端的最末5个核苷酸内。例如,在此类实施方案中,对于23个核苷酸的RNAi剂,与C9orf72基因互补区域的链通常不含位于中心13个核苷酸处的任何错配。本文中描述的方法或本领域中已知的方法可用以确定,含有与靶标序列的错配的RNAi剂在抑制C9orf72基因的表达中是否有效。考虑具有错配的RNAi剂在抑制C9orf72基因的表达中的效力是重要的,尤其是C9orf72基因中的特定互补区域是已知具有该群内的多态性序列变化时。

[0239] 如本文中所用,“实质上全部核苷酸是经修饰的”是指大多数并非全部经修饰,且

可包括不超过5、4、3、2或1个未经修饰的核苷酸。

[0240] 如本文中所用，术语“正义链”或“过客链(passenger strand)”是指RNAi剂的链，其包括与如本文中定义术语的反义链的区域实质上互补的区域。

[0241] 如本文中所用，术语“裂解区域”是指位于紧邻裂解位点处的区域。裂解位点是靶标上的裂解出现处的位点。在一些实施方案中，裂解区域包含位于裂解位点任一端且紧邻该裂解位点的三个碱基。在一些实施方案中，裂解区域包含位于裂解位点任一端且紧邻该裂解位点的两个碱基。在一些实施方案中，裂解位点特异性地出现于反义链的由核苷酸10及11为边界的位点处，且裂解区域包含核苷酸11、12及13。

[0242] 如本文中所用且除非明确排除，否则当术语“互补”用来描述关于第二核苷酸序列的第一核苷酸序列时，是指包含寡核苷酸或多核苷酸的第一核苷酸序列在某些条件下与包含寡核苷酸或多核苷酸的第二核苷酸序列杂交及形成双链体结构的能力，如普通技术人员所理解的。例如，这些条件可以是严格条件，其中严格条件可包括:400mM NaCl、40mM PIPES pH 6.4、1mM EDTA、50℃或70℃、12至16小时，之后洗涤(参见，例如，《分子克隆:实验室手册》(“Molecular Cloning:A Laboratory Manual,Sambrook,et al.(1989)Cold Spring Harbor Laboratory Press))。可施加其他条件，诸如生理学相关条件如可在有机体内部遭遇的。普通技术人员将能够根据所杂交的核苷酸的最终应用而确定最适用于两个序列的互补性测试的条件集。

[0243] 如本文中所述的RNAi剂例如dsRNA内的互补序列包括包含寡核苷酸或多核苷酸的第一核苷酸序列与包含寡核苷酸或多核苷酸的第二核苷酸序列在一个或两个核苷酸序列的整体长度上的碱基配对。此类序列可称为彼此“完全互补”。然而，本文中，若第一序列称为与第二序列“基本上互补”，则当杂交形成至多30个碱基对的双链体时，两个序列可以是完全互补，或它们可形成一个或多个但通常不超过5、4、3或2个错配碱基对，同时保留在最适于其最终应用的条件下杂交的能力，如对经由RISC途径的基因表达的抑制。然而，若两个寡核苷酸设计为当杂交时形成一个或多个单链突出端，则此类突出端不应视为关于确定互补性的错配。例如，包含一个长度为21个核苷酸的寡核苷酸及另一个长度为23个核苷酸的寡核苷酸的dsRNA，其中该较长的核苷酸包含一个与该较短的核苷酸完全互补的21个核苷酸的序列，对于本文所描述的目的，仍可称为“完全互补”。

[0244] 如本文中所用，“互补的”序列还可包括非Watson-Crick碱基对或从非天然核苷酸及经修饰的核苷酸形成的碱基对，或完全由其形成，只要对其杂交能力的上述需求得以满足即可。此类非Watson-Crick碱基对包括但不限于，G:U Wobble碱基配对或Hoogstein碱基配对。

[0245] 本文中，术语“互补”、“完全互补”及“实质上互补”可关于dsRNA的正义链与反义链之间或RNAi剂的反义链与靶标序列之间的碱基配对而使用，如可从其用途的语境中理解的。

[0246] 如本文中所用，与信使RNA(mRNA)的“至少一部分实质上互补”的多核苷酸是指与感兴趣的mRNA(例如，编码C9orf72的mRNA)的连续部分实质上互补的多核苷酸。例如，如果该序列与编码C9orf72的mRNA的非中断部分实质上互补，则该多核苷酸与C9orf72 mRNA的至少一部分互补。

[0247] 据此，在一些实施方案中，本文中披露的反义多核苷酸与靶标C9orf72序列完全互

补。在其他实施方案中，本文披露的反义多核苷酸与靶标C9orf72序列实质上互补，且包含连续核苷酸序列，该连续核苷酸序列在其整个长度上与SEQ ID NO:121至124及133中任一者的核苷酸序列的等效区域(equivalent region)为至少80％互补，诸如约85％、约90％、约％91％、约92％、约93％、约94％、约95％、约96％、约97％、约98％、或约99％互补。

[0248] 在一些实施方案中，本文披露的反义多核苷酸与靶标C9orf72序列的片段实质上互补，并且包含连续核苷酸序列，该连续核苷酸序列在其整个长度上与选自由SEQ ID NO:121的核苷酸1594至1616、802至824、239至261、1308至1330、233至255、1595至1617、240至262、1532至1554、237至259、3268至3290、806至828、1620至1642、526至548、1169至1191、1266至1288、1247至1269、586至608、1257至1279、及400至422所组成的组的SEQ ID NO:121的片段为至少约80％互补，诸如约85％、约90％、约91％、约92％、约93％、约94％、约95％、约96％、约95％、约97％、约98％或约99％互补。

[0249] 在一些实施方案中，本文披露的反义多核苷酸与靶标C9orf72序列的片段实质上互补，并且包含连续核苷酸序列，该连续核苷酸序列在其整个长度上与SEQ ID NO:133的片段(诸如SEQ ID NO:133的核苷酸200至290，例如，核苷酸200至290、200至280、200至270、200至260、200至250、200至240、200至230、200至225、210至290、210至280、210至270、210至260、210至250、210至240、210至235、220至290、220至280、220至270、220至260、220至250、220至245、225至250、225至245、230至290、230至280、230至270、230至260、230至255、235至260、240至265、240至290、240至280、240至270、240至265、250至290、250至280、250至275、260至290、260至285)为至少80％互补，诸如约85％、约90％、约91％、约92％、约93％、约94％、约95％、约96％、约97％、约98％、或约99％互补。上文列举的范围之间的范围也视为本公开的一部分。

[0250] 在其他实施方案中，本文披露的反义多核苷酸与靶标C9orf72序列实质上互补，并且包含连续核苷酸序列，该连续核苷酸序列是在其整个长度上与表10A、10B、12至15、19或20任一者中的正义链核苷酸序列的任一个、或表10A、10B、12至15、19或20任一者中的正义链核苷酸序列的任一个的片段为至少约80％互补，诸如约85％、约90％、约91％、约92％、约93％、约94％、约95％、约96％、约95％、约97％、约98％、约99％、或100％互补。

[0251] 在一个实施方案中，本公开的RNAi剂包括正义链，该正义链与反义多核苷酸实质上互补，该反义多核苷酸进而与靶标C9orf72序列相同，其中该正义链多核苷酸包含连续核苷酸序列，该连续核苷酸序列是在其整个长度上与SEQ ID NO:125至128及134的核苷酸序列的等效区域为至少约80％互补，诸如约85％、约90％、约91％、约92％、约93％、约94％、约95％、约96％、约95％、约97％、约98％、约99％、或100％互补。

[0252] 在一些实施方案中，本发明的iRNA包含正义链，该正义链与反义多核苷酸实质上互补，该反义多核苷酸进而与靶标C9orf72序列互补，其中该正义链多核苷酸包含连续核苷酸序列，该连续核苷酸序列是在其整个长度上与表10A、10B、12至15、19或20任一者中的反义链核苷酸序列的任一个、或表10A、10B、12至25、19或20任一者中的反义链核苷酸序列的任一个的片段为至少约80％互补，诸如约85％、约90％、约91％、约92％、约93％、约94％、约95％、约96％、约95％、约97％、约98％、约99％、或100％互补。

[0253] 在某些实施方案中，该正义链及反义链选自下列双链体的任一者：AD-348904.1、AD-348136.1、AD-347612.1、AD-348639.1、AD-347606.1、AD-348905.1、AD-347613.1、AD-

348842.1、AD-347610.1、AD-350329.1、AD-348140.1、AD-348930.1、AD-347863.1、AD-348500.1、AD-348597.1、AD-348578.1、AD-347923.1、AD-348588.1、AD-347773.1及AD-348136.1。

[0254] 在某些实施方案中，该正义链及反义链选自下列双链体的任一者：AD-463863.1、AD-463862.1、AD-463869.1、AD-463873.1、AD-463872.1及AD-463860.1。

[0255] 在一个实施方案中，对C9orf72基因的至少部分的抑制是借由C9orf72 mRNA(例如，正义mRNA、反义mRNA、总C9orf72 mRNA、正义含C9orf72重复的mRNA和/或反义含C9orf72重复的RNA)量的降低而评估，其是从第一细胞或细胞的群组中分离或检测，在该细胞或细胞的群组中，C9orf72基因被转录，且该细胞或细胞的群组已经处理，使得与实质上与该第一细胞或细胞的群组相同但尚未经所述处理的第二细胞或细胞的群组(对照细胞)相比，C9orf72基因的表达得以抑制。抑制的程度可根据以下表达：

[0256] $$\frac{(\text{对照细胞中的mRNA}) - (\text{经处理细胞中的mRNA})}{(\text{对照细胞中的mRNA})} \bullet 100\%$$

[0257] 如本文中所用，短语“使细胞与RNAi剂接触”(例如dsRNA)包括借由任何可能的手段接触细胞。使细胞与RNAi剂接触包括在体外使细胞与RNAi剂接触或在体内使细胞与RNAi剂接触。接触可直接或间接进行。因此，例如，RNAi剂可通过单独执行该方法而使其与细胞物理接触，或作为另一种选择，可将RNAi剂置于将允许或造成其后续与该细胞接触的情况。

[0258] 例如，可通过使用RNAi剂温育细胞而使该细胞在体外接触该RNAi剂。例如，可通过将RNAi剂注射至细胞所处的组织内或邻近该组织处，或通过将RNAi剂注射至另一区域例如中枢神经系统(CNS)，任选的经由鞘内注射、玻璃体腔内注射或其他注射，或注射至血流或皮下空间内，使得该剂将后续到达待接触的细胞所处的组织，从而使该细胞在体内与该RNAi剂接触。例如，RNAi剂可含有配体或与其偶联，该配体是例如下文描述的一个或多个亲脂性部分并进一步详述于例如PCT/US2019/031170中，该专利通过援引并入本文，该配体引导或以其他方式稳定化位于感兴趣的位点例如CNS处的RNAi剂，比如体外接触方法与体内接触方法的组合也是可能的。例如，细胞可在体外与RNAi剂接触，并随后移植入受试者体内。

[0259] 在一个实施方案中，使细胞与RNAi剂接触包括，通过促进或影响至该细胞内的摄取或吸收而“将RNAi剂引入或递送至细胞内”。RNAi剂的吸收或摄取可通过无辅助的扩散或活性细胞过程而进行，或通过辅助剂或装置而进行。将RNAi剂引入细胞内可以是体外或体内过程。例如，对于体内进行的引入，可将RNAi剂注射至组织部位或系统性给药。在体外进行的引入细胞内包括本领域中已知的方法，如电穿孔及脂质转染。其他途径描述于下文中或是本领域中已知的。

[0260] 术语“亲脂体”或“亲脂性部分”广义上是指具有对于脂质的亲和性的任何化合物或化学部分。一种表征亲脂性部分的亲脂性的方法是通过辛醇-水分配系数$\log K_{ow}$进行，其中K_{ow}是两相系统处于平衡状态时，化学品在辛醇相中的浓度与其在水相中的浓度的比率。辛醇-水分配系数是物质的实验室测量的特性。然而，其还可使用归因于化学品的结构组分的系数进行预测，这些系数是使用第一原理或经验方法计算(参见，例如，Tetko et al., J.Chem.Inf.Comput.Sci.41:1407-21(2001)，其通过援引以其整体并入本文)。其提供物质

倾向于非水性或油性介质而非水(即,其亲水/亲脂平衡)的热力学测量。原则上,当化学物质的$logK_{ow}$超过0时,该物质是亲脂性的。典型地,亲脂性部分具备超过1、超过1.5、超过2、长3、超过4、超过5或超过10的$logK_{ow}$。例如,6-氨基己醇的$logK_{ow}$是例如预测为大约0.7。使用同一方法,胆固醇基N-(己-6-醇)氨基甲酸酯的$logK_{ow}$预测为10.7。

[0261] 分子的亲脂性可随其携带的官能基团而改变。例如,将羟基或胺基团加至亲脂性部分的末端,可增加或降低该亲脂性部分的分配系数(例如,$logK_{ow}$)值。

[0262] 另选地,与一个或多个亲脂性部分缀合的双链RNAi剂的疏水性可通过其蛋白质结合特征而测量。例如,在某些实施方案中,双链RNAi剂的血浆蛋白结合测定中的未结合部分可确定为与该双链RNAi剂的相对疏水性正相关,而相对疏水性可与该双链RNAi剂的沉默活性正相关。

[0263] 在一个实施方案中,血浆蛋白结合测定是使用人血清白蛋白蛋白质的电泳迁移率变动分析测定。这一结合测定的示例性方案在例如PCT/US2019/031170中详细例证。通过该结合测定中的未结合siRNA部分测量的双链RNAi剂的疏水性超过0.15、超过0.2、超过0.25、超过0.3、超过0.35、超过0.4、超过0.45或超过0.5,以增强siRNA的体内递送。

[0264] 据此,将亲脂性部分缀合至双链RNAi剂的内部位置,提供优化的疏水性,以增强siRNA的体内递送。

[0265] 术语“脂质纳米颗粒”或“LNP”是媒介,其包含脂质层,该脂质层封装药学活性分子如核酸分子例如RNAi剂或RNAi剂自其转录的脂质体。LNP描述于,例如,美国专利第6,858,225号、第6,815,432号、第8,158,601号及第8,058,069号,这些专利的全部内容通过援引而并入本文。

[0266] 如本文所用,“受试者”是动物诸如哺乳动物,包括灵长动物(诸如人,非人灵长动物例如猴及黑猩猩)或非灵长动物(诸如大鼠或小鼠)。在优选实施方案中,受试者是人,如正在进行对将受益于C9orf72表达降低的疾病、疾患或病症的治疗或评估的人;处于将受益于靶标C9orf72 RNA水平降低的疾病、疾患或病症的风险下的人;患有将受益于C9orf72表达降低的疾病、疾患或病症的人;或正在进行对将受益于C9orf72表达降低的疾病、疾患或病症的治疗的人,如本文中所述。在一些实施方案中,受试者是女人。在一些实施方案中,受试者是男人。在一个实施方案中,受试者是成年受试者。在一个实施方案中,受试者是小儿受试者。在另一实施方案中,受试者是青少年受试者,即,年龄小于20岁的受试者。

[0267] 如本文所用,术语“治疗”或“治疗”是指有益或所希望的结果,包括但不限于,一种或多种与C9orf72基因表达或C9orf72蛋白产生相关(例如C9orf72相关疾病,诸如C9orf72相关疾病)的体征或症状的缓解或减轻。“治疗”还可意指相对于在治疗缺失情况下预期的生存期而延长生存期。

[0268] 在受试者的C9orf72或疾病标记物或症状水平的情境中,术语“降低”是指此水平的统计学显著的降低。该降低可以是,例如,至少10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%或更多。在某些实施方案中,降低是至少20%。在某些实施方案中,该降低是疾病标记物降低至少50%,例如,含有正义的团簇或含有反义的团簇的水平和/或异常二肽重复蛋白质的水平的降低,例如,降低50%、55%、60%、65%、70%、75%、80%、85%、90%、95%或更多。在一些实施方案中,疾病标记物(例如C9orf72蛋白质和/或表达水平)降低不超过50%,例如,不超过50%、45%、

40％、35％、30％、25％、20％、15％、10％或5％。在受试者体内的C9orf72水平的情境中，“降低”优选是低至如同不具此病变的个体正常范围内所接受的水平。在某些实施方案中，“降低”是患有疾病的受试者的标记物或症状水平与处于个体正常范围内所接受水平之间的差异的下降，例如，肥胖个体的体重与具有正常范围内所接受重量的个体的体重之间的下降水平。

[0269] 如本文所用，当关于将受益于C9orf72基因表达或C9orf72蛋白产生的降低的疾病、疾患或其病症而使用“预防”或“阻止”时，该术语是指降低受试者将发展出与此疾病、疾患或病症相关的症状例如C9orf72相关疾病的症状的可能性。没有发展出疾病、疾患或症状，或减小与此疾病、疾患或症状相关的症状的发展(例如，将该疾病或疾患的临床可接受规模减小至少约10％)，或显现症状的延迟(例如，延迟几天、几周、几个月或几年)，是视为有效的预防。

[0270] 如本文所用，术语“C9orf72相关疾病”或“C9orf72相关疾患”包括将会受益于C9orf72表达和/或活性的降低的任何疾病或疾患。示例性C9orf72相关疾病包括其中受试者在C9orf72基因中位于外显子1a与1b之间的内含子中携带六核苷酸重复序列(GGGCC)扩增的那些疾病，例如，肌萎缩侧索硬化(ALS)或额颞叶痴呆(FTD)以及由于C9orf72扩增引起的亨廷顿样综合征。

[0271] 正常的G4C2重复是约25个单位或更少，而高渗透性疾病等位基因典型是大于约60个重复单位，范围最高超过4,000个单位；罕见地，47至60个单位之间的重复与家族性疾病区分开。重复引物PCR测定典型用来检测较小的扩增(<80)，但准确地确定较大重复的尺寸需要其他技术(例如，Southern blot杂交)，其提供对于长度的估算。

[0272] 在C9orf72基因的内含子中具有GGGGCC(或G4C2)六核苷酸扩增的受试者，即使在同一家族中也可表达为肌萎缩侧索硬化(ALS)或额颞叶痴呆(FTD)，因此，与此扩增相关的神经退行在本文中称为“C9orf72肌萎缩侧索硬化/额颞叶痴呆”或“C9orf72 ALS/FTD”。其是常染色体显性遗传病并且是最常见的家族性ALS形式，占家族性ALS的约三分的一以及ALS门诊散发病例的5％至10％。它也是FTD的常见病因，造成家族性FTD的四分之一。症状发作的年龄从30岁到70岁，平均在五十几岁的后期发作。C9orf72介导的ALS最经常与典型ALS相似，可以是延髓发作或四肢发作，可进展迅速(虽然不总是如此)，并且可能与后来的认知症状相关联。因此，C9orf72介导的ALS的评估及治疗与在任何ALS患者中进行的一样。C9orf72介导的FTD的最常见模式是行为变异性FTD，伴有全范围的行为及认知症状，包括去抑制、冷漠及执行功能障碍。较少见地，C9orf72介导的FTD表现出语义变异性原发性进行性失语症(semantic variant primary progressive aphasia(PPA))或不流利变异性PPA，并且非常罕见地，可能与皮质基底综合征、进行性核上神经麻痹症或HD样综合征相似。偶尔，帕金森特征可见于C9orf72介导的ALS或FTD中。

[0273] 受试者可以展现出以行为的进行性改变、执行功能障碍和/或语言障碍为特征的额颞叶退行(frontotemporal lobar degeneration,FTLD)。三种FTLD临床症状中，行为变异性FTD(bvFTD)最常见，但并非唯一。其以进行性行为障碍及执行功能的持续减退为特征，并在脑MRI上出现明显的额颞叶萎缩。还可表达为运动神经元疾病，包括上或下运动神经元功能障碍(或两者)，可能符合或不符合完全ALS表型标准。某种程度的帕金森综合征存在于多个患有C9orf72相关bvFTD的个体中，是典型的无震颤的强直类型，并且对左旋多巴无响

应。

[0274] 亨廷顿样综合征(HD样综合征或HDL综合征)属于遗传性神经退行性疾病家族,其在下列方面与亨廷顿氏病(HD)极为相似,它们典型产生舞蹈症、认知减退或痴呆以及行为或精神问题的组合。

[0275] 患有由于C9orf72扩增引起的亨廷顿样综合征的受试者被表征为具有行动障碍,包括张力失常、舞蹈症、肌阵挛、震颤及强直。相关特征还有认知及记忆障碍、早期精神紊乱及行为问题。平均发作年龄为约43岁(范围8至60岁)。早期精神及行为问题(包括抑郁、冷漠、强迫性行为及思觉失调)是常见的。认知综合征表达为执行功能障碍。行动障碍是突出的:还可能表现为帕金森综合征特征及锥体特征(pyramidal feature)。如本文中所用,"治疗有效量"旨在包括RNAi剂的下述的量,当将该RNAi剂给药至患有C9orf72相关疾病的受试者时,其量足以有效治疗该疾病(例如,通过削弱、缓解或维持现有疾病或疾病的一种或多种症状)。"治疗有效量"可依据RNAi剂、该剂如何给药、疾病及其严重性及待治疗的受试者的病史、年龄、体重、家族病史、基因组成、先前治疗或并行治疗的类型(若存在)、以及其他个体特征而变。

[0276] 如本文中所用,"预防有效量"旨在包括RNAi剂的下述的量,当将该RNAi剂给药至患有C9orf72相关疾患的受试者时,其量足以预防或缓解该疾病或该疾病的一种或多种症状。缓解该疾病包括减缓该疾病的进程或减轻后来发展的疾病的严重性。"预防有效量"可依据RNAi剂、该剂如何给药、疾病风险程度、及待治疗的患者的病史、年龄、体重、家族病史、基因组成、先前治疗或并行治疗的类型(若存在)、以及其他个体特征而变。

[0277] "治疗有效量"或"预防有效量"还包括以可用于任何治疗的合理效益/风险比率产生一些所希望的局部或系统性效果的RNAi剂的量。本公开的方法中采用的RNAi剂可以足以产生可用于此治疗的合理效益/风险比率的量给药。

[0278] 本文中,短语"药学上可接受"用以是指那些化合物、材料、组合物或剂型,其处于适用于与人类受试者及动物受试者的组织接触而无过度毒性、刺激、过敏反应或其他问题或并发症的安全医疗判断的范畴内,与合理效益/风险比率相称。

[0279] 如本文中所用,短语"药学上可接受的载剂"意指药学上可接受的材料、组合物或媒介,如液体或固体填料、稀释剂、赋形剂、制造助剂(例如,润滑剂、滑石、硬脂酸镁、硬脂酸钙、硬脂酸锌、或硬脂酸)、或溶剂封装材料,其涉及将受试化合物从一个器官或身体部分携带或运送至另一器官或身体部分。就与制剂的其他成分兼容且不损害被治疗的受试者而言,每一载剂必须是"可接受"的。可用作药学上可接受的载剂的材料的一些示例包括:(1)糖类,如乳糖、葡萄糖及蔗糖;(2)淀粉类,如玉米淀粉及马铃薯淀粉;(3)纤维素及其衍生物,如羧甲基纤维素钠、乙基纤维素及醋酸纤维素;(4)黄蓍胶粉末;(5)麦芽;(6)明胶;(7)润滑剂,如硬脂酸镁、十二烷基硫酸钠及滑石;(8)赋形剂,如可可脂及栓剂蜡;(9)油类,如花生油、棉籽油、葵花籽油、芝麻油、橄榄油及大豆油;(10)二醇类,如丙二醇;(11)多元醇类,如甘油、山梨醇、甘露醇及聚乙二醇;(12)酯类,如油酸乙酯及月桂酸乙酯;(13)琼脂;(14)缓冲剂,如氢氧化镁及氢氧化铝;(15)海藻酸;(16)无热原水;(17)等张盐水;(18)林格氏溶液;(19)乙醇;(20)pH缓冲溶液;(21)聚酯类、聚碳酸酯类或聚酐类;(22)增量剂(bulking agent),如多肽及氨基酸类;(23)血清组分,如血清白蛋白、HDL及LDL;以及(22)药学制剂中采用的其他非毒性兼容物质。

[0280] 如本文中所用，术语“样本”包括从受试者分离的相似体液、细胞或组织的集合，以及受试者体内存在的体液、细胞或组织。生物体液的示例包括血液、血清及浆膜液、血浆、脑脊液、眼液、淋巴液、尿液、唾液等。组织样本可包括来自组织、器官或局部区域的样本。例如，样本可源自特定的器官、器官的部分、或那些器官内的体液或细胞。在某些实施方案中，样本可源自脑(例如，全脑或脑的某些节段，例如，纹状体，或脑内的某些类型的细胞，诸如神经元或胶质细胞(星状细胞、寡树突细胞、微胶质细胞))。在一些实施方案中，“源自受试者的样本”是指从该受试者抽取的血液或自其衍生的血浆或血清。在进一步的实施方案中，“源自受试者的样本”是指源自受试者的脑组织(或其亚组分)或视网膜组织(或其亚组分)。

[0281] II.本公开的RNAi剂

[0282] 如本文他处所描述，C9orf72中的突变已经与家族性额颞叶痴呆(FTD)及肌萎缩侧索硬化(ALS)关联。该突变是定位在C9orf72基因的外显子1A与外显子1B间的G4C2(SEQ ID NO:1)六核苷酸重复扩增的结果。六核苷酸重复序列可经由非AUG启动的机制翻译。含重复扩增的RNA(靶标RNA)的累积或重复序列的翻译可造成或有促进FTD和/或ALS或与FTD和/或ALS相关的疾病症状。

[0283] 据此，本发明提供dsRNA RNAi剂，其选择性且有效地降低与该六核苷酸重复扩增相关联的C9orf72相关表达产物、RNA和/或经翻译的多肽的表达。在一些实施方案中，dsRNA剂靶向(例如，选择性地靶向)含六核苷酸重复的RNA(靶标RNA)，并且敲低该靶标RNA及自该含六核苷酸重复的RNA表达的多肽。dsRNA剂可用于对与FTD和/或ALS相关的体征或症状进行治疗性治疗和/或预防的方法中，体征或症状包括但不限于，RNA团簇的重复序列长度依赖性形成、特异性RNA结合蛋白质的隔离、以及神经元中重复序列相关非AUG(AUG)翻译所致的二肽重复蛋白质(例如，聚(甘氨酸-丙氨酸)、聚(甘氨酸-脯氨酸)、聚(甘氨酸-精氨酸)、聚(丙氨酸-脯氨酸)及聚(脯氨酸-精氨酸))的累积及集聚。dsRNA剂可用于对FTD和/或ALS相关的体征或症状进行治疗性治疗和/或预防的方法中，体征或症状包括但不限于，运动神经元疾病的体征或症状以及痴呆的体征或症状。运动神经元疾病的体征及症状包括，例如，绊倒、掉落物品、胳膊和/或腿的不正常疲劳、口齿不清、肌肉痉挛及抽搐、笑或哭不受控制以及呼吸困难。痴呆的体征及症状可包括，例如，行为改变、性格改变、言语及语言问题及行动相关问题。此类方法包括将一种或多种本文所描述的dsRNA剂给药至受试者(例如，人或动物受试者)。

[0284] 本文所描述的dsRNA剂可终止或降低含重复的C9orf72 RNA(例如，作为RNA团簇而测定)的累积，从而阻止通过RNA翻译进行的二肽重复蛋白质的合成。

[0285] 在一些实施方案中，本发明的dsRNA剂靶向成熟C9orf72 mRNA(即，其内含子已经剔除的mRNA)。在其他实施方案中，本发明的dsRNA剂靶向含有内含子诸如内含子1A的C9orf72 RNA(即，其内含子尚未剔除的mRNA、剔除前体mRNA的RNA区域或另选经剪接的RNA)。

[0286] 在一些实施方案中，本发明的dsRNA剂靶向C9orf72转录本，该转录本在重复上游的非编码外显子1A处起始。C9orf72重复扩增是显性突变，其基本上总是作为杂合突变而发生。因此，在一些实施方案中，本文所描述的dsRNA剂靶向在外显子1A处起始的转录本，并且主要仅影响致病性等位基因；正常的等位基因几乎唯一地在重复序列下游的外显子1B处启动转录。因此，正常(非致病性)转录本得以保留，故该dsRNA剂不会消灭所有C9orf72蛋白质

的产生。

[0287] 在一个实施方案中,RNAi剂包括用于抑制细胞中C9orf72基因表达的双链核糖核酸(dsRNA)分子,该细胞诸如处于受试者例如哺乳动物诸如患有C9orf72相关疾病例如C9orf72相关疾病的人的体内。该dsRNA包括具有互补区域的反义链,该互补区域与在C9orf72基因表达中所形成的mRNA的至少一部分互补。互补区域为约15至30个核苷酸的长度或更短。当与表达C9orf72基因的细胞接触时,RNAi剂将该C9orf72基因(例如,人类基因、灵长动物基因、非灵长动物基因)的表达抑制至少50%,如通过例如PCR或基于分支DNA(bDNA)的方法所分析,或通过基于蛋白质的方法所分析,例如通过使用例如western blot的免疫萦光分子或流式细胞术所分析。在优选的实施方案中,敲低的水平是于人神经胚细胞瘤BE(2)C细胞中使用下文实施例1中提供的双重萦光素酶测定法测定。

[0288] 在一些实施方案中,本文所描述的dsRNA剂设计为用以终止或降低含重复的C9orf72 RNA(例如,作为RNA团簇而测定)的累积,从而阻止通过RNA翻译进行的二肽重复蛋白质的合成。为达成此目标,在某些实施方案中,dsRNA剂靶向C9orf72 mRNA蛋白质编码序列(参见,例如,表12及表13)。在其他实施方案中,dsRNA剂靶向C9orf72转录本,该转录本在重复的上游的非编码外显子1A处起始。在重复的下游起始的转录本将不含致病性RNA。C9orf72重复序列扩增是显性突变,其基本上总是作为杂合突变而发生。因此,当本文所描述的dsRNA剂靶向在外显子1A处起始的转录本时,它们主要仅影响致病性等位基因;正常的等位基因几乎只在重复的下游的外显子1B处启动转录,因此,正常(非致病性)转录本得以保留。

[0289] dsRNA包括两个RNA链,在该dsRNA将被使用的条件下,这两条链互补并杂交以形成双链体结构。dsRNA的一条链(反义链)包括互补区域,该互补区域与靶标序列实质上互补,且通常完全互补,以靶向序列。靶标序列可源自在C9orf72基因表达过程中形成的mRNA序列。另一条链(正义链)包括如下区域,该区域与该反义链互补,使得当在适宜条件下组合时,两条链杂交并形成双链体结构。如本文中他处所述,dsRNA的互补序列还可作为单独核酸分子的自互补区域而包含,与作为单独寡核苷酸相反。

[0290] 通常,该双链体结构是15至30个碱基对的长度,如15至29、15至28、15至27、15至26、15至25、15至24、15至23、15至22、15至21、15至20、15至19、15至18、15至17、18至30、18至29、18至28、18至27、18至26、18至25、18至24、18至23、18至22、18至21、18至20、19至30、19至29、19至28、19至27、19至26、19至25、19至24、19至23、19至22、19至21、19至20、20至30、20至29、20至28、20至27、20至26、20至25、20至24、20至23、20至22、20至21、21至30、21至29、21至28、21至27、21至26、21至25、21至24、21至23、或21至22个碱基对的长度。在某些优选实施方案中,双链体结构是18至25个碱基对的长度,例如,18至25、18至24、18至23、18至22、18至21、18至20、19至25、19至24、19至23、19至22、19至21、19至20、20至25、20至24、20至23、20至22、20至21、21至25、21至24、21至23、21至22、22至25、22至24、22至23、23至25、23至24或24至25个碱基对的长度,例如,19至21个碱基对的长度。上文列举的范围及长度之间的范围及长度也视为本公开的一部分。

[0291] 同样,与靶标序列互补的区域是15至30个核苷酸的长度,例如,15至29、15至28、15至27、15至26、15至25、15至24、15至23、15至22、15至21、15至20、15至19、15至18、15至17、18至30、18至29、18至28、18至27、18至26、18至25、18至24、18至23、18至22、18至21、18至20、19

至30、19至29、19至28、19至27、19至26、19至25、19至24、19至23、19至22、19至21、19至20、20至30、20至29、20至28、20至27、20至26、20至25、20至24、20至23、20至22、20至21、21至30、21至29、21至28、21至27、21至26、21至25、21至24、21至23、或21至22个核苷酸的长度，例如，19至23个核苷酸的长度或21至23个核苷酸的长度。上文列举的范围及长度之间的范围及长度也视为本公开的一部分。

[0292] 在一些实施方案中，该双链体结构是19至30个碱基对的长度。同样，与靶标序列互补的区域是19至30个核苷酸的长度。

[0293] 在一些实施方案中，dsRNA是15至23个核苷酸的长度，19至23个核苷酸的长度，或25至30个核苷酸的长度。通常，该dsRNA足够长，以用作切丁酶的底物。例如，本领域中熟知，长度超过约21至23个核苷酸的dsRNA可用作切丁酶的底物。本领域普通技术人员还应意识到，裂解靶向的RNA区域最通常是较大RNA分子的一部分，一般为mRNA分子。若相关，则mRNA靶标的“一部分”是mRNA靶标的连续序列，其长度足以使其作为RNAi引导的裂解(即，通过RISC途径裂解)的底物。

[0294] 本领域普通技术人员还应意识到，双链体区域是dsRNA的主要官能部分，例如，双链体区域是约15至36个碱基对，例如，15至36、15至35、15至34、15至33、15至32、15至31、15至30、15至29、15至28、15至27、15至26、15至25、15至24、15至23、15至22、15至21、15至20、15至19、15至18、15至17、18至30、18至29、18至28、18至27、18至26、18至25、18至24、18至23、18至22、18至21、18至20、19至30、19至29、19至28、19至27、19至26、19至25、19至24、19至23、19至22、19至21、19至20、20至30、20至29、20至28、20至27、20至26、20至25、20至24、20至23、20至22、20至21、21至30、21至29、21至28、21至27、21至26、21至25、21至24、21至23、或21至22个碱基对，例如，19至21个碱基对。因此，在一个实施方案中，就其被加工为(例如15至30个碱基对的)靶向期望的RNA以用于裂解的的官能性双链体的程度而言，具有超过30个碱基对的双链区域的RNA分子或RNA分子的复合物是dsRNA。因此，本领域普通技术人员应意识到，在一个实施方案中，miRNA是dsRNA。在另一实施方案中，dsRNA不是天然出现的miRNA。在另一实施方案中，可用于靶向C9orf72表达的RNAi剂并非通过较大dsRNA的裂解而在靶标细胞内生成的。

[0295] 本文中所述的dsRNA可进一步包括一个或多个具有例如1、2、3或4个核苷酸的单链核苷酸突出端。核苷酸突出端可包含核苷酸/核苷类似物或由其组成，其中该核苷酸/核苷类似物包括脱氧核苷酸/核苷。所述突出端可位于正义链、反义链或其任何组合上。此外，突出端的核苷酸可存在于dsRNA的反义链或正义链的5’端、3’端或两端。

[0296] dsRNA可通过本领域中已知的标准方法合成。本发明的双链RNAi化合物可使用两步的过程制备。首先，双链RNA分子的单个链是单独制备的。随后，将这些组分链退火。该siRNA化合物的单个链可使用溶液相有机合成、固相有机合成或两者制备。有机合成提供下述优点：包含非天然或经修饰的核苷酸的寡核苷酸链可容易制备。同样，本发明的单链寡核苷酸可使用溶液相有机合成、固相有机合成或两者制备。

[0297] 在某些实施方案中，本发明的dsRNA剂靶向C9orf72靶标RNA，该靶标RNA包含六核苷酸重复，该重复包含SEQ ID NO:1的多个连续拷贝，例如，C9orf72靶向具有致病性六核苷酸重复扩增(具有，例如，至少约30、至少约35、至少约40、至少约50、至少约60、至少约70、至少约80、至少约100、至少约200、至少约300、至少约400或至少约500个六核苷酸重复)的

RNA。

[0298] 本文所描述的dsRNA剂可靶向含六核苷酸重复的C9orf72 RNA(C9orf72靶标RNA)的任何区域。在某些实施方案中,本文所描述的dsRNA可靶向位于C9orf72 RNA的外显子1A起始位点与外显子2起始位点之间(即,包括外显子1A但不包括外显子2)的核苷酸序列(即,对应于从相对应的C9orf72基因中的外显子1A起始位点到外显子2起始位点的区域)。在某些实施方案中,本文所描述的dsRNA剂可靶向位于C9orf72 RNA的外显子1A起始位点与外显子1B终止位点之间(即,包括外显子1A及外显子1B)的核苷酸序列。在某些实施方案中,本文所描述的dsRNA剂可靶向位于C9orf72 RNA的外显子1A起始位点与外显子1B起始位点之间(即,包括外显子1A但不包括外显子1B)的核苷酸序列。在某些实施方案中,本文所描述的dsRNA剂可靶向位于C9orf72 RNA的外显子1A起始位点与六核苷酸重复序列终止位点之间(即,包括外显子1A及六核苷酸重复序列)的核苷酸序列。在某些实施方案中,本文所描述的dsRNA剂可靶向位于C9orf72 RNA的外显子1A起始位点与六核苷酸重复序列起始位点之间(即,包括外显子1A但不包括六核苷酸重复序列)的核苷酸序列。在某些实施方案中,本文所描述的dsRNA剂可靶向位于C9orf72 RNA的外显子1A起始位点与外显子1A终止位点之间的核苷酸序列。

[0299] 在某些实施方案中,本文所描述的dsRNA剂可靶向位于C9orf72RNA的外显子1A终止位点与外显子2起始位点之间(即,不包括外显子1A且不包括外显子2)的核苷酸序列。在某些实施方案中,本文所描述的dsRNA剂可靶向位于C9orf72 RNA的外显子1A终止位点与外显子1B终止位点之间(即,不包括外显子1A但包括外显子1B)的核苷酸序列。在某些实施方案中,本文所描述的dsRNA剂可靶向位于C9orf72 RNA的外显子1A终止位点与外显子1B起始位点之间(即,不包括外显子1A且不包括外显子1B)的核苷酸序列。在某些实施方案中,本文所描述的dsRNA剂可靶向位于C9orf72 RNA的外显子1A终止位点与六核苷酸重复序列终止位点之间(即,不包括外显子1A但包括六核苷酸重复序列)的核苷酸序列。在某些实施方案中,本文所描述的dsRNA剂可靶向位于C9orf72 RNA的外显子1A终止位点与六核苷酸重复序列起始位点之间(即,不包括外显子1A且不包括六核苷酸重复序列)的核苷酸序列。

[0300] 在某些实施方案中,本文所描述的dsRNA剂可靶向位于C9orf72RNA的六核苷酸重复序列起始位点与外显子2起始位点之间(即,包括六核苷酸重复序列但不包括外显子2)的核苷酸序列。在某些实施方案中,本文所描述的dsRNA剂可靶向位于C9orf72 RNA的六核苷酸重复序列起始位点与外显子1B终止位点之间(即,包括六核苷酸重复序列及外显子1B)的核苷酸序列。在某些实施方案中,本文所描述的dsRNA剂可靶向位于C9orf72 RNA的六核苷酸重复序列起始位点与外显子1B起始位点之间(即,包括六核苷酸重复序列但不包括外显子1B)的核苷酸序列。在某些实施方案中,本文所描述的dsRNA剂可靶向位于C9orf72 RNA的六核苷酸重复序列起始位点与六核苷酸重复序列终止位点之间的核苷酸序列。

[0301] 在某些实施方案中,本文所描述的dsRNA剂可靶向位于C9orf72RNA的六核苷酸重复序列终止位点与外显子2起始位点之间(即,不包括六核苷酸重复序列且不包括外显子2)的核苷酸序列。在某些实施方案中,本文所描述的dsRNA剂可靶向位于C9orf72 RNA的六核苷酸重复序列终止位点与外显子1B终止位点之间(即,不包括六核苷酸重复序列但包括外显子1B)的核苷酸序列。在某些实施方案中,本文所描述的dsRNA剂可靶向位于C9orf72 RNA的六核苷酸重复序列终止位点与外显子1B起始位点之间(即,不包括六核苷酸重复序列且

不包括外显子1B)的核苷酸序列。

[0302] 在某些实施方案中,本文所描述的dsRNA剂可靶向位于C9orf72RNA的外显子1B起始位点与外显子2起始位点之间(即,包括外显子1B且不包括外显子2)的核苷酸序列。

[0303] 在某些实施方案中,本文所描述的dsRNA剂可靶向位于C9orf72RNA的外显子1B起始位点与外显子1B终止位点之间的核苷酸序列。

[0304] 在某些实施方案中,本文所描述的dsRNA剂可靶向位于C9orf72RNA的外显子1B终止位点与外显子2起始位点之间(即,不包括外显子1B且不包括外显子2)的核苷酸序列。

[0305] 一些dsRNA剂靶向C9orf72 RNA的内含子区域。例如,本文所描述的dsRNA剂可靶向位于C9orf72基因的外显子1A与外显子1B之间的内含子(即,位于C9orf72靶标RNA的外显子1A终止位点与外显子1B起始位点之间的内含子)。dsRNA剂所靶向的区域可处于C9orf72靶标RNA的外显子1A与六核苷酸重复序列间的内含子序列中该六核苷酸重复序列的上游(5')或下游(3')。在具体实施例中,dsRNA剂靶向位于C9orf72靶标RNA中六核苷酸重复序列的上游(5')的区域。例如,dsRNA剂可靶向位于外显子1A终止位点与六核苷酸重复序列起始位点之间的内含子序列。

[0306] 在一个实施方案中,dsRNA剂靶向C9orf72靶标RNA中的区域,该区域对应于人类C9orf72基因的位于人类基因组组件GRCh38/hg38坐标chr9:27,567,165-27,573,866之间的区域(对应于从外显子1A起始位点到外显子2起始位点(即,位于外显子2之前的内含子终止位点)的区域;从这一区域转录的具有三个GGGGCC重复的例示性RNA序列详述于SEQ ID NO:111中,且反向互补序列详述于SEQ ID NO:112中)。在另一具体实施例中,dsRNA剂靶向C9orf72靶标RNA中的区域,该区域对应于人类C9orf72基因的位于人类基因组组件GRCh38/hg38坐标chr9:27,573,494-27,573,708之间的区域(对应于外显子1A与外显子1B之间(即,位于外显子1A终止位点与外显子1B起始位点之间)的内含子序列的区域;从这一区域转录的具有三个GGGGCC重复的例示性RNA序列详述于SEQ ID NO:113中,且反向互补序列详述于SEQ ID NO:114中)。在另一具体实施例中,dsRNA剂靶向C9orf72靶标RNA中的区域,该区域对应于人类C9orf72基因的位于人类基因组组件GRCh38/hg38坐标chr9:27,573,547-27,573,708之间的区域(对应于外显子1A与六核苷酸重复序列起始位点之间(即,位于外显子1A终止位点与六核苷酸重复序列起始位点之间)的内含子序列的区域;从这一区域转录的例示性RNA序列详述于SEQ ID NO:115中,且反向互补序列详述于SEQ ID NO:116中)。在另一具体实施例中,dsRNA剂靶向C9orf72靶标RNA中的区域,该区域对应于人类C9orf72基因的位于人类基因组组件GRCh38/hg38坐标chr9:27,573,605-27,573,640之间的区域(对应于来自位于外显子1A与六核苷酸重复序列起始位点之间的内含子区域区域的片段;从这一区域转录的例示性RNA序列详述于SEQ ID NO:117中,且反向互补序列详述于SEQ ID NO:118中)。

[0307] 在某些实施方案中,本发明的dsRNA剂靶向SEQ ID NO:133的核苷酸200至290。

[0308] 在某些实施方案中,dsRNA剂靶向来自下列的任一核苷酸序列:SEQ ID NO:121的核苷酸230至270、233至262、800至840、800至830、802至828、1240至1290、1240至1280、1247至1288、1590至1645、1590至1620及1594至1642。

[0309] 在某些实施方案中,dsRNA剂靶向来自下列的任一核苷酸序列:SEQ ID NO:121的核苷酸1594至1616、802至824、239至261、1308至1330、233至255、1595至1617、240至262、

1532至1554、237至259、3268至3290、806至828、1620至1642、526至548、1169至1191、1266至1288、1247至1269、586至608、1257至1279及400至422。

[0310] 一方面，本公开的dsRNA包括至少两个核苷酸序列：正义序列及反义序列。C9orf72的正义链序列可选自表10A、表10B、表12至15、表19或表20的任一者中提供的序列所组成的组，而该正义链的反义链的相应核苷酸序列可选自表10A、表10B、表12至15、表19或表20中任一者的序列所组成的组。在此方面，这两个序列中的一个与这两个序列的另一者互补，且这些序列中的一个与在C9orf72基因的表达中生成的mRNA序列实质上互补。由此，在此方面，dsRNA将包括两个寡核苷酸，其中一个寡核苷酸描述为表10A、表10B、表12至表15、表19或表20中任一者的正义链(过客链)，而第二个寡核苷酸描述为表10A、表10B、表12至表15、表19或表20中任一者的正义链的对应反义链(导引链)。

[0311] 在一个实施方案中，该dsRNA的实质上互补序列包含在独立的寡核苷酸上。在另一实施方案中，该dsRNA的实质上互补序列包含在单寡核苷酸上。

[0312] 应理解，尽管表10A、表10B、表12至表15、表19或表20中的序列描述为经修饰或缀合的序列，但本公开的RNAi剂的RNA，例如，本公开的dsRNA，可包含于表10A、表10B、表12至表15、表19或表20中任一者中详述的任一序列，其是未修饰的、未缀合的或不同于其中所描述的的经修饰或缀合的。例如，尽管表13及15中所示的剂的正义链缀合至GalNAc配体，但这些剂可缀合至引导CNS递送的部分，例如，C16配体，如本文所描述。亲脂性配体可包括于本申请中提供的任何位置处。

[0313] 本领域技术人员应知悉，具有约20至23个碱基对如21个碱基对的双链体结构的dsRNA已经被认为在诱导RNA干扰中尤其有效(Elbashir et al.,(2001)EMBO J.,20:6877-6888)。然而，其他人已经发现，更短或更长的RNA双链体结构也可以是有效的(Chu and Rana(2007)RNA 14:1714-1719;Kim et al.(2005)Nat Biotech 23:222-226)。在上文描述的实施方案中，由于本文中提供的寡核苷酸序列的性质，本文所描述的dsRNA可包括至少一条长度为最少21个核苷酸的链。可合理地预期，仅在一端或两端减去几个核苷酸的较短双链体可能具有与上述dsRNA类似的效果。因此，具有源自本文所提供的一个序列的至少15、16、17、18、19、20或更多个连续核苷酸的序列，且使用以例如如本文实施例中提供的Be(2)c细胞及10nM浓度的RNA剂进行的体外测定以及PCR测定所测的其抑制C9orf72基因表达的能力与包含全序列dsRNA的能力相差不超过约5%、10%、15%、20%、25%或30%的dsRNA，是预期处于本公开的范畴内。

[0314] 此外，本文所描述的RNA鉴定C9orf72转录本中易受RISC介导的裂解影响的位点。由此，本公开进一步提出位于此位点内的RNAi剂。如本文中所用，如果RNAi剂促进该特定位点内任意处的转录本的裂解，则称该RNAi剂为以RNA转录本的特定位点内为靶向。此RNAi剂通常包括来自本文所提供的一个序列的至少约15个连续核苷酸，优选至少19个核苷酸，该连续核苷酸与来自C9orf72基因中所选序列的连续区域的另一核苷酸序列偶联。

[0315] 本文披露的dsRNA剂抑制包含六核苷酸重复的C9orf72靶标RNA的表达。抑制表达包括任何水平的抑制(例如，对表达的部分抑制)。例如，dsRNA剂可将包含六核苷酸重复的C9orf72靶标RNA的表达抑制至少约10%、至少约20%、至少约30%、至少约40%、至少约50%、至少约60%、至少约70%、至少约80%、或至少约90%(或抑制到C9orf72靶标RNA无法检测的点)。例如，给药至表达包含六核苷酸重复的C9orf72靶标RNA的细胞之后，在24至48

小时内，达到这些抑制水平。降低可以是，例如，相对于以dsRNA剂处理之前的细胞或相对于未经dsRNA剂处理的对照细胞。

[0316] 本文所描述的dsRNA剂还可，例如，相对于成熟C9orf72信使RNA的表达而选择性地抑制包含内含子六核苷酸重复的C9orf72靶标RNA的表达。本语境中，成熟C9orf72信使RNA是已经剪接及加工的C9orf72 RNA转录本。成熟C9orf72信使RNA唯一地由外显子组成并且内含子全部已经移除。在将dsRNA剂给药至表达C9orf72靶标RNA的细胞后，如果C9orf72靶标RNA的表达的降低大于成熟C9orf72信使RNA的表达的相对降低，则dsRNA剂可相对于成熟C9orf72信使RNA的表达而选择性地抑制包含内含子六核苷酸重复的C9orf72靶标RNA的表达。例如，dsRNA剂可将成熟C9orf72信使RNA的表达抑制少于约50％、少于约40％、少于约30％、少于约20％、少于约10％或少于约5％(或者，例如，对于表达不具有任何统计学显著或功能上显著的效应)。例如，给药至表达包含成熟C9orf72信使RNA的细胞之后，在24至48小时内，达到这些抑制水平。

[0317] 本文所披露的dsRNA剂还可例如降低细胞中二肽重复蛋白质合成或二肽重复蛋白质水平(例如，在给药至细胞后24至48小时)。例如，dsRNA剂可将二肽重复蛋白质合成或二肽重复蛋白质水平降低至少约10％、至少约20％、至少约30％、至少约40％、至少约50％、至少约60％、至少约70％、至少约80％、或至少约90％。降低可以是，例如，相对于以dsRNA剂处理之前的细胞或相对于未经dsRNA剂处理的对照细胞。

[0318] III.本公开的经修饰的RNAi剂

[0319] 在一个实施方案中，本公开的RNAi剂的RNA例如dsRNA是未修饰者，且不包含例如本领域中已知及本文所述的化学修饰或缀合。在优选实施方案中，本公开的RNAi剂的RNA(例如dsRNA)经化学修饰以增强稳定性或其他有益特征。在本公开的某些实施方案中，本公开的RNAi剂的实质上全部核苷酸是经修饰。在本公开的其他实施方案中，本公开的RNAi剂的全部核苷酸是经修饰。本公开的“实质上全部核苷酸是经修饰的”是指大多数并非全部经修饰，且可包括不超过5、4、3、2或1个未经修饰的核苷酸。在本公开的又其他实施方案中，本公开的RNAi剂可包括不超过5、4、3、2或1个经修饰的核苷酸。

[0320] 本公开提出的核酸可通过本领域中良好构建的方法合成或修饰，该方法是例如那些于《现代核酸化学技术》(“Current protocols in nucleic acid chemistry”，Beaucage,S.L.et al.(Edrs.),John Wiley&Sons,Inc.,New York,NY,USA)中描述的，该文献通过援引而并入本文。修饰包括，例如，末端修饰，例如，5’末端修饰(磷酸化、缀合、反向键联)或3’末端修饰(缀合、DNA核苷酸、反向键联等)；碱基修饰，例如，替换为稳定化碱基、去稳定化碱基、或与同伴的拓展物进行碱基配对的碱基，移除碱基(无碱基的核苷酸)，或缀合碱基；糖修饰(例如，在2’-位置或4’-位置)或糖的取代；或骨架修饰，包括磷酸二酯键的修饰或取代。可用于本文所述实施方案中的RNAi剂的具体示例包括，但不限于，含有经修饰的骨架或不含天然核苷酸间键的RNA。具有经修饰的骨架的RNA除此之外还包括那些在骨架中不具有磷原子的。对于本说明书的目的，且如本领域中有时参照的，在其核苷酸间骨架中不具有磷原子的经修饰的RNA也可视为寡核苷酸。在一些实施方案中，经修饰的RNAi剂将在其核苷酸间骨架中具有磷原子。

[0321] 经修饰的RNA骨架包括，例如，硫代磷酸酯、手性硫代磷酸酯、二硫代磷酸酯、磷酸三酯、氨基烷基磷酸三酯、包括3’-亚烷基磷酸酯及手性磷酸酯的甲基及其他烷基磷酸酯、

次膦酸酯、包括3'-氨基磷酸酯及氨基烷基氨基磷酸酯的氨基磷酸酯、硫羰基氨基磷酸酯、硫羰基烷基磷酸酯、硫羰基烷基磷酸三酯、以及具有正常3'-5'键的硼磷酸酯，这些的2'-5'键联类似物，以及具有反向极性的那些，其中相邻的核苷单元对是将3'-5'键联至5'-3'或将2'-5'键联至5'-2'。还可包括多种盐类、混合盐类及游离酸形式。在本发明的一些实施方案中，本发明的dsRNA剂可以是游离酸形式。在本发明的其他实施方案中，本发明的dsRNA剂可以是盐形式。在一个实施方案中，本发明的dsRNA剂可以是钠盐形式。在某些实施方案中，当本发明的dsRNA剂是钠盐形式时，钠离子可作为该剂中存在的实质上全部磷酸二酯和/或硫代磷酸酯基团的抗衡离子而存在于该剂中。其中实质上全部磷酸二酯和/或硫代磷酸酯键都具有钠抗衡离子的剂，包括不超过5、4、3、2或1个没有钠抗衡离子的磷酸二酯和/或硫代磷酸酯键。在一些实施方案中，当本发明的dsRNA剂是钠盐形式时，钠离子可作为该剂中存在的全部磷酸二酯和/或硫代磷酸酯基团的抗衡离子而存在于该剂中。

[0322] 教导上述含磷键的制备的代表性美国专利包括但不限于，美国专利第3,687,808号、第4,469,863号、第4,476,301号、第5,023,243号、第5,177,195号、第5,188,897号、第5,264,423号、第5,276,019号、第5,278,302号、第5,286,717号、第5,321,131号、第5,399,676号、第5,405,939号、第5,453,496号、第5,455,233号、第5,466,677号、第5,476,925号、第5,519,126号、第5,536,821号、第5,541,316号、第5,550,111号、第5,563,253号、第5,571,799号、第5,587,361号、第5,625,050号、第6,028,188号、第6,124,445号、第6,160,109号、第6,169,170号、第6,172,209号、第6,239,265号、第6,277,603号、第6,326,199号、第6,346,614号、第6,444,423号、第6,531,590号、第6,534,639号、第6,608,035号、第6,683,167号、第6,858,715号、第6,867,294号、第6,878,805号、第7,015,315号、第7,041,816号、第7,273,933号、第7,321,029号、及美国再公告专利第39464号，其各自的全部内容通过援引而并入本文。

[0323] 其内部不包括磷原子的经修饰的RNA骨架具有通过短链烷基或环烷基核苷酸间键、混合杂原子及烷基或环烷基核苷酸间键、或一个或多个短链杂原子或杂环核苷酸间键形成的骨架。这些包括那些具有吗啉代键(部分地由核苷的糖部分形成)；硅氧烷骨架；硫醚、亚砜及砜骨架；甲酰基及硫代甲酰基骨架；亚甲基甲酰基及硫代甲酰基骨架；含有亚烷基的骨架；氨基磺酸酯骨架；亚甲基亚氨基及亚甲基肼基骨架；磺酸酯及磺酰胺骨架；酰胺骨架；以及其他具有混合的N、O、S及CH_2组分部分的。

[0324] 教导上述寡核苷酸的制备的代表性美国专利包括但不限于，美国专利第5,034,506号、第5,166,315号、第5,185,444号、第5,214,134号、第5,216,141号、第5,235,033号、第5,64,562号、第5,264,564号、第5,405,938号、第5,434,257号、第5,466,677号、第5,470,967号、第5,489,677号、第5,541,307号、第5,561,225号、第5,596,086号、第5,602,240号、第5,608,046号、第5,610,289号、第5,618,704号、第5,623,070号、第5,663,312号、第5,633,360号、第5,677,437号、及第5,677,439号，其各自的全部内容通过援引而并入本文。

[0325] 在其他实施方案中，适宜的RNA模拟物预期用于RNAi剂中，其中该核苷酸单元的糖及核苷酸间键两者(即骨架)被置换有新颖基团。碱基单元维持与适宜的核酸靶标化合物杂交。一种此类寡聚化合物，已经显示具有优异杂交特性的RNA模拟物，称为肽核酸(PNA)。在PNA化合物中，RNA的糖骨架替换为含有酰胺的骨架，尤其是氨基乙基甘油骨架。核碱基得以保留，且直接或间接地键联至骨架的酰胺部分的氮杂氮原子。教导PNA化合物的制备的代表

性美国专利包括但不限于,美国专利第5,539,082号、第5,714,331号、及第5,719,262号,其各自的全部内容通过援引而并入本文。适用于本公开的RNAi剂中的额外的PNA化合物描述于,例如,Nielsen et al.,Science,1991,254,1497-1500中。

[0326] 本发明提出的一些实施方案包括具有硫代磷酸酯骨架的RNA以及具有杂原子骨架的寡核苷酸,尤其是上文引用的美国专利第5,489,677号的--CH_2--NH--CH_2-、--CH_2--N(CH_3)--O--CH_2--[称为亚甲基(甲基亚胺基)或MMI骨架]、--CH_2--O--N(CH_3)--CH_2--、--CH_2--N(CH_3)--N(CH_3)--CH_2--及--N(CH_3)--CH_2--CH_2--[其中,天然磷酸二酯骨架表示为--O--P--O--CH_2--],以及上文引用的美国专利第5,602,240号的酰胺骨架。在一些实施方案中,本文提出的RNA具有上文引用的美国专利第US 5,034,506号的吗啉代骨架结构。

[0327] 经修饰的RNA还可含有一个或多个经取代的糖部分。本文提出的RNAi剂例如dsRNA可包括位于2'位置的下述中的一个:OH;F;O-、S-或N-烷基;O-、S-或N-烯基;O-、S-或N-炔基;或O-烷基-O-烷基,其中该烷基、烯基及炔基可以是经取代或未经取代的C_1至C_{10}烷基或C_2至C_{10}烯基及炔基。示例性的合适修饰包括O[(CH_2)$_n$O]$_m$$CH_3$、O($CH_2$).$_n$$OCH_3$、O($CH_2$)$_n$$NH_2$,O($CH_2$)$_n$$CH_3$、O($CH_2$)$_nONH_2$、及O($CH_2$)$_n$ON[($CH_2$)$_n$$CH_3$)]$_2$,其中,n及m是1至约10。在其他实施方案中,dsRNA包括位于2'位置的下述中的一个:C_1至C_{10}低级烷基、经取代的低级烷基、烷芳基、芳烷基、O-烷芳基或O-芳烷基、SH、SCH_3、OCN、Cl、Br、CN、CF_3、OCF_3、$SOCH_3$、SO_2CH_3、ONO_2、NO_2、N_3、NH_2、杂环烷基、杂环烷基芳基、氨基烷基氨基、聚烷基氨基、经取代的硅烷基、RNA裂解基团、保护基团、嵌入剂、用于改善RNAi的药物动力学特性的基团、或用于改善RNAi的药效动力学特性的基团、以及其他具有类似特性的取代基。在一些实施方案中,该修饰是包括2'甲氧基乙氧基(2'-O--$CH_2CH_2OCH_3$,也称为2'-O-(2-甲氧基乙基)或2'-MOE)(Martin et al.,Helv.Chim.Acta,1995,78:486-504),即,烷氧基-烷氧基。另一例示性修饰是2'-二甲基氨基氧代乙氧基,即,O(CH_2)$_2$ON(CH_3)$_2$基团,也称为2'-DMAOE,如下文实施例中所述;以及2'-二甲基氨基乙氧基乙氧基(本领域中也称为2'-O-二甲基氨基乙氧基乙基或2'-DMAEOE),即,2'-O--CH_2--O--CH_2--N(CH_2)$_2$。其他示例性修饰包括:5'-Me-2'-F核苷酸、5'-Me-2'-OMe核苷酸、5'-Me-2'-脱氧核苷酸(在这三组中,都是R异构体与S异构体两者);2'-烷氧基烷基;以及2'-NMA(N-甲基乙酰胺)。

[0328] 其他修饰包括2'-甲氧基(2'-OCH_3)、2'-氨基丙氧基(2'-$OCH_2CH_2CH_2NH_2$)、2'-O-十六烷基及2'-氟(2'-F)。类似的修饰还可在RNAi剂的RNA的其他位置作成,尤其是3'末端核苷酸的糖的3'位置或2'-5'键的dsRNA中以及5'末端核苷酸的5'位置。RNAi剂还可具有替代呋喃戊糖基糖的糖模拟物如环丁基部分。教导此类经修饰的糖结构的制备的代表性美国专利包括但不限于,美国专利第4,981,957号、第5,118,800号、第5,319,080号、第5,359,044号、第5,393,878号、第5,446,137号、第5,466,786号、第5,514,785号、第5,519,134号、第5,567,811号、第5,576,427号、第5,591,722号、第5,597,909号、第5,610,300号、第5,627,053号、第5,639,873号、第5,646,265号、第5,658,873号、第5,670,633号、及第5,700,920号,这些中的某些为本申请所共有。前述者各自的全部内容通过援引并入本文。

[0329] 本公开的RNAi剂还可包括核碱基(本领域中一般简称为"碱基")修饰或取代。如本文中所用,"未经修饰"或"天然"核碱基包括嘌呤碱基腺嘌呤(A)及鸟嘌呤(G),以及嘧啶碱基胸腺嘧啶(T)、胞嘧啶(C)及尿嘧啶(U)。经修饰的核碱基包括其他合成及天然核碱基,如5-甲基胞嘧啶(5-me-C);5-羟甲基胞嘧啶;黄嘌呤;次黄嘌呤;2-氨基腺嘌呤;腺嘌呤及鸟嘌

呤的6-甲基及其他烷基衍生物;腺嘌呤及鸟嘌呤的2-丙基及其他烷基衍生物;2-硫尿嘧啶、2-硫胸腺嘧啶、2-硫胞嘧啶;5-卤尿嘧啶、5-卤胞嘧啶;5-丙炔基尿嘧啶、5-丙炔基胞嘧啶;6-偶氮尿嘧啶、6-偶氮胞嘧啶、6-偶氮胸腺嘧啶;5-尿嘧啶(假尿嘧啶);4-硫尿嘧啶;8-卤、8-氨基、8-巯基、8-硫烷基、8-羟基及其他8-取代的腺嘌呤及鸟嘌呤;5-卤尤其是5-溴、5-三氟甲基及其他5-取代的尿嘧啶及胞嘧啶;7-甲基鸟嘌呤及7-甲基腺嘌呤;8-氮杂鸟嘌呤及8-氮杂腺嘌呤;7-去氮鸟嘌呤及7-去氮腺嘌呤;以及3-去氮鸟嘌呤及3-去氮腺嘌呤。进一步的核碱基包括那些披露于美国专利第3,687,808号中的;那些披露于《生物化学、生物技术及医药中的经修饰的核苷酸》(Modified Nucleosides in Biochemistry,Biotechnology and Medicine,Herdewijn,P.ed.Wiley-VCH,2008)中的;那些披露于《聚合物科学及工程简明百科》(The Concise Encyclopedia Of Polymer Science And Engineering,pages 858-859,Kroschwitz,J.L,ed.John Wiley&Sons,1990)中的;那些由Englisch et al.,(1991)Angewandte Chemie,International Edition 30:613披露的;以及那些由《dsRNA研究及应用》第15章第289至302页(Sanghvi,Y S.,Chapter 15,dsRNA Research and Applications,pages 289-302,Crooke,S.T.and Lebleu,B.,Ed.,CRC Press,1993)披露的。这些核碱基中的某些尤其可用于增加本公开提出的寡聚化合物的结合亲和性。这些包括5-取代的嘧啶、6-氮杂嘧啶及N-2、N-6及O-6取代的嘌呤,包括2-氨基丙基腺嘌呤、5-丙基尿嘧啶及5-丙基胞嘧啶。5-甲基胞嘧啶取代已经显示将核酸双链体稳定性增加0.6至1.2℃(Sanghvi,Y.S.,Crooke,S.T.and Lebleu,B.,Eds.,dsRNA Research and Applications,CRC Press,Boca Raton,1993,pp.276-278)且是示例性的碱基取代,尤其是当与2'-O-甲氧基乙基糖修饰组合时。

[0330] 教导某些上述经修饰的核碱基以及其他经修饰的核碱基的代表性美国专利包括但不限于,上述的美国专利第3,687,808号、第4,845,205号、第5,130,30号、第5,134,066号、第5,175,273号、第5,367,066号、第5,432,272号、第5,457,187号、第5,459,255号、第5,484,908号、第5,502,177号、第5,525,711号、第5,552,540号、第5,587,469号、第5,594,121号、第5,596,091号、第5,614,617号、第5,681,941号、第5,750,692号、第6,015,886号、第6,147,200号、第6,166,197号、第6,222,025号、第6,235,887号、第6,380,368号、第6,528,640号、第6,639,062号、第6,617,438号、第7,045,610号、第7,427,672号、及第7,495,088号,其各自的全部内容通过援引而并入本文。

[0331] 本公开的RNAi剂还可经修饰,以包括一个或多个锁核酸(LNA)。锁核酸是具有经修饰的核糖部分的核苷酸,其中该核糖部分包含连接2'碳与4'碳的外接桥。这一结构有效地将该核糖"锁定"为3'-环内结构的构象。将锁核酸加至siRNA中已经显示增加血清中siRNA稳定性,且降低脱靶效应(Elmen,J.et al.,(2005)Nucleic Acids Research 33(1):439-447;Mook,OR.et al.,(2007)Mol Canc Ther 6(3):833-843;Grunweller,A.et al.,(2003)Nucleic Acids Research 31(12):3185-3193)。

[0332] 本公开的RNAi剂还可经修饰,以包括一个或多个双环糖部分。"双环糖"是通过两个原子的桥接而修饰的呋喃糖基环。"双环核苷"("BNA")是核苷,其具有包含连接该糖环的两个碳原子的桥的糖部分,从而形成双环系统。在某些实施方案中,桥连接糖环的4'碳与2'碳。因此,在一些实施方案中,本公开的剂可包括一个或多个锁核酸(LNA)。锁核酸是具有经修饰的核糖部分的核苷酸,其中该核糖部分包含连接2'碳与4'碳的外接桥。换言之,LNA是

包含双环糖部分的核苷酸，其中该双环糖部分包含4'-CH_2-O-2'桥。这一结构有效地将该核糖"锁定"为3'-环内结构的构象。将锁核酸加至siRNA中已经显示增加血清中siRNA稳定性，且降低脱靶效应(Elmen,J.et al.,(2005)Nucleic Acids Research 33(1):439-447;Mook,OR.et al.,(2007)Mol Canc Ther 6(3):833-843;Grunweller,A.et al.,(2003)Nucleic Acids Research 31(12):3185-3193)。用于本公开的多核苷酸的双环核苷的示例包括而不限于，包含位于4'核糖基环原子与2'核糖基环原子间的桥的核苷。在某些实施方案中，本公开的反义多核苷酸剂包括一个或多个包含4'至2'桥的双环核苷。此类4'至2'桥接的双环核苷的示例包括但不限于，4'-(CH2)—O-2'(LNA)；4'-(CH2)—S-2'；4'-(CH2)2—O-2'(ENA)；4'-CH(CH3)—O-2'(也称为"受约束的乙基"或"cEt")及4'-CH(CH2OCH3)—O-2'(及其类似物；参见，例如，美国专利第7,399,845号)；4'-C(CH3)(CH3)—O-2'(及其类似物；参见，例如，美国专利第8,278,283号)；4'-CH2—N(OCH3)-2'(及其类似物；参见，例如，美国专利第8,278,425号)；4'-CH2—O—N(CH3)-2'(参见，例如，美国专利申请第2004/0171570号)；4'-CH2—N(R)—O-2'，其中R是H、C1-C12烷基、或保护基团(参见，例如，美国专利第7,427,672号)；4'-CH2—C(H)(CH3)-2'(参见，例如，Chattopadhyayaet al.,J.Org.Chem.,2009,74,118-134)；即4'-CH2—C(=CH2)-2'(及其类似物；参见，例如，美国专利第8,278,426号)。前述者各自的全部内容通过援引并入本文。

[0333] 教导锁核酸核苷酸的制备的其他代表性美国专利及美国专利公开包括但不限于下列者：美国专利第6,268,490号、第6,525,191号、第6,670,461号、第6,770,748号、第6,794,499号、第6,998,484号、第7,053,207号、第7,034,133号、第7,084,125号、第7,399,845号、第7,427,672号、第7,569,686号、第7,741,457号、第8,022,193号、第8,030,467号、第8,278,425号、第8,278,426号、第8,278,283号、第2008/0039618号、及第2009/0012281号，其各自的全部内容通过援引而并入本文。

[0334] 前述双环核苷中的任一者可制备为具有一种或多种立体化学糖构型，包括，例如，α-L-呋喃核糖及β-D-呋喃核糖(参见，WO 99/14226)。

[0335] 本公开的RNAi剂还可经修饰，以包括一个或多个约束的乙基核苷酸。如本文中所用，"约束的乙基核苷酸"或"cEt"是包含双环糖部分的锁核酸，其中该双环糖部分包含4'-CH(CH3)-O-2'桥。在一个实施方案中，约束的乙基核苷酸是S构象，本文中称为"S-cEt"。

[0336] 本公开的RNAi剂还可经修饰，以包括一个或多个"构象上受限的核苷酸"("CRN")。CRN是具有连接核糖的C2'碳与C4'碳或核糖的C3碳与C5'碳的接头的核苷酸类似物。CRN将该核糖锁定为适宜的构象，且增加其与mRNA的杂交亲和性。该接头足够长，以将氧置于对于稳定性及亲和性为最优的位置，从而使核糖不易折叠(puckering)。

[0337] 教导上述CRN的制备的代表性专利公开包括但不限于，US 2013/0190383及WO 2013/036868，其各自的全部内容通过援引而并入本文。

[0338] 在一些实施方案中，本公开的RNAi剂包含一个或多个作为UNA(未锁核酸)核苷酸的单体。UNA是未锁定的非环状核酸，其中该糖的任何键已经移除，形成未锁定的"糖"残基。在一示例中，UNA还涵盖其C1'-C4'键(即，位于C1'碳与C4'碳间的碳-氧-碳共价键)已经移除的单体。在另一示例中，糖的C2'-C3'键(即，位于C2'碳与C3'碳之间的碳-碳共价键)已经移除(参见，Nuc.Acids Symp.Series,52,133-134(2008)及Fluiter et al.,Mol.Biosyst.,2009,10,1039，通过援引并入本文)。

[0339] 教导UNA的制备的代表性美国专利公开包括但不限于,US 8,314,227及美国专利公开第2013/0096289号、第2013/0011922号、及第2011/03130200号,其各自的全部内容通过援引而并入本文。

[0340] 对RNA分子的端的潜在稳定化修饰可包括N-(乙酰基氨基己酰基)-4-羟基脯氨醇(Hyp-C6-NHAc)、N-(己酰基-4-羟基脯氨醇(Hyp-C6)、N-(乙酰基-4-羟基脯氨醇(Hyp-NHAc)、胸腺嘧啶-2'-O-脱氧胸腺嘧啶(醚)、N-(氨基己酰基)-4-羟基脯氨醇(Hyp-C6-氨基)、2-二十二酰基-尿苷-3″-磷酸酯、反向碱基dT(idT)等。这一修饰的披露可见于WO 2011/005861中。

[0341] 本公开的RNAi剂的其他修饰包括5'磷酸酯或5'磷酸酯模拟物,例如位于RNAi剂的反义链上的5'端磷酸酯或磷酸酯模拟物。合适的磷酸酯模拟物披露于,例如,美国专利公开第2012/0157511号中,其全部内容通过援引而并入本文。

[0342] A.本公开的包含基序的经修饰的RNAi剂

[0343] 在本公开的某些方面,本公开的双链RNAi剂包括具有如例如WO 2013/075035中所披露的化学修饰的剂,该申请的全部内容通过援引而并入本文。如本文及WO 2013/075035中所示,通过将一个或多个位于三个连续核苷酸上的三个相同修饰的基序引入RNAi剂的正义链或反义链中,尤其在裂解位点处或邻近裂解位点处,可获得杰出的结果。在一些实施方案中,RNAi剂的正义链及反义链可以其他方式完全修饰。这些基序的引入中断正义链或反义链的修饰模式(若存在)。RNAi剂可任选的与亲脂性配体例如C16配体例如于正义链缀合。RNAi剂可任选的经(S)-二醇核酸(GNA)修饰例如在反义链的一个或多个残基上修饰。所得RNAi剂呈现杰出的基因沉默活性。

[0344] 据此,本公开提供能在体内抑制靶标基因(即,C9orf72基因)的表达的双链RNAi剂。RNAi剂包含正义链及反义链。RNAi剂的每一条链可以是15至30个核苷酸的长度。例如,每一条链可以是16至30个核苷酸的长度、17至30个核苷酸的长度、25至30个核苷酸的长度、27至30个核苷酸的长度、17至21个核苷酸的长度、17至19个核苷酸的长度、19至25个核苷酸的长度、19至23个核苷酸的长度、19至21个核苷酸的长度、21至25个核苷酸的长度、或21至23个核苷酸的长度。在某些实施方案中,每一条链是19至23个核苷酸的长度。

[0345] 该正义链及反义链典型形成双链体双链RNA("dsRNA"),本文中还称为"RNAi剂"。RNAi剂的双链体区域可以是15至30个核苷酸对的长度。例如,该双链体区域可以是16至30个核苷酸对的长度、17至30个核苷酸对的长度、27至30个核苷酸对的长度、17至23个核苷酸对的长度、17至21个核苷酸对的长度、17至19个核苷酸对的长度、19至25个核苷酸对的长度、19至23个核苷酸对的长度、19至21个核苷酸对的长度、21至25个核苷酸对的长度、或21至23个核苷酸对的长度。在另一示例中,该双链体区域是选自15、16、17、18、19、20、21、22、23、24、25、26及27个核苷酸的长度。在优选的实施方案中,该双链体区域是19至21个核苷酸的长度。

[0346] 在一个实施方案中,该RNAi剂可含有位于每条链或两条链的3'端、5'端或两端的一个或多个突出端区域或封端基团。突出端可以为1至6个核苷酸的长度,例如,2至6个核苷酸的长度、1至5个核苷酸的长度、2至5个核苷酸的长度、1至4个核苷酸的长度、2至4个核苷酸的长度、1至3个核苷酸的长度、2至3个核苷酸的长度、或1至2个核苷酸的长度。在优选的实施方案中,该核苷酸突出端区域是2个核苷酸的长度。突出端可以是一条链比另一条链长

的结果，或是相同长度的两条链交错的结果。突出端可形成与靶标mRNA的错配，或其可与待作为靶标的基因序列互补或可以是另一序列。该第一链与第二链还可通过例如额外的碱基连接以形成发夹来连接，或通过其他非碱基接头来连接。

[0347] 在一个实施方案中，该RNAi剂的突出端区域中的核苷酸可各自独立为经修饰或未经修饰的核苷酸，包括但不限于，2'-糖修饰，诸如，2-F、2'-O-甲基胸苷(T)及其任何组合。

[0348] 例如，TT可以是任一链任一端的突出端序列。突出端可形成与靶标mRNA的错配，或其可与待作为靶标的基因序列互补或可以是另一序列。

[0349] 位于该RNAi剂的正义链、反义链或两个的5'突出端或3'突出端可经磷酸化。在一些实施方案中，突出端区域含有两个核苷酸且在这两个核苷酸之间具有硫代磷酸酯，其中，两个核苷酸可以是相同或不同。在一个实施方案中，该突出端是存在于正义链、反义链或两条链的3'端。在一个实施方案中，这一3'突出端存在于反义链中。在一个实施方案中，这一3'突出端存在于正义链中。

[0350] 该RNAi剂可仅含有单个突出端，该突出端可强化该RNAi剂的干扰活性而不影响其整体稳定性。例如，该单链突出端可位于正义链的3'末端，或者位于反义链的3'末端。RNAi剂还可具有钝端，位于反义链的5'末端(或正义链的3'末端)，反之亦然。通常，该RNAi的反义链具有位于3'末端的核苷酸突出端，且5'端是钝端。尽管不希望受理论束缚，但位于反义链5'末端的不对称钝端以及反义链的3'末端突出端有助于将导引链加载至RISC过程中。

[0351] 在一个实施方案中，该RNAi剂是19个核苷酸长度的双钝端，其中正义链含有至少一个位于从5'端计数第7、8、9位置处的三个连续核苷酸上的三个2'-F修饰的基序。反义链含有至少一个位于从5'端计数第11、12、13位置处的三个连续核苷酸上的三个2'-O-甲基修饰的基序。

[0352] 在另一实施方案中，该RNAi剂是20个核苷酸长度的双钝端，其中正义链含有至少一个位于从5'端计数第8、9、10位置处的三个连续核苷酸上的三个2'-F修饰的基序。反义链含有至少一个位于从5'端计数第11、12、13位置处的三个连续核苷酸上的三个2'-O-甲基修饰的基序。

[0353] 在又一实施方案中，该RNAi剂是21个核苷酸长度的双钝端，其中正义链含有至少一个位于从5'端计数第9、10、11位置处的三个连续核苷酸的三个2'-F修饰的基序。反义链含有至少一个位于从5'端计数第11、12、13位置处的三个连续核苷酸的三个2'-O-甲基修饰的基序。

[0354] 在一个实施方案中，该RNAi剂包含21个核苷酸的正义链及23个核苷酸的反义链，其中该正义链含有至少一个位于从5'端计数第9、10、11位置处的三个连续核苷酸上的三个2'-F修饰的基序；该反义链含有至少一个位于从5'端计数第11、12、13位置处的三个连续核苷酸上的三个2'-O-甲基修饰的基序，其中该RNAi剂的一端是钝端而另一端包含具有2个核苷酸的突出端。优选地，具有2个核苷酸的突出端位于反义链的3'端。当具有2个核苷酸的突出端位于反义链的3'端时，在末端三个核苷酸之间可能存在两个硫代硫酸酯核苷酸间键，其中，该三个核苷酸中的两者是突出端核苷酸，且第三个核苷酸与紧邻该突出端核苷酸的下一个核苷酸配对。在一个实施方案中，RNAi剂在正义链的5'端及反义链的5'端两处额外具有位于末端三个核苷酸之间的两个硫代磷酸酯核苷酸间键。在一个实施方案中，该RNAi剂的正义链及反义链中的每一个核苷酸，包括作为这些基序的一部分的核苷酸，都是经修

饰的核苷酸。在一个实施方案中,每一残基是独立经2'-O-甲基或3'-氟以例如交替基序方式修饰。任选的,RNAi剂进一步包含配体(例如,亲脂性配体,任选的C16配体)。

[0355] 在一个实施方案中,RNAi剂包含正义链及反义链,其中该正义链是25至30个核苷酸残基的长度,其中第一链的从5'末端核苷酸(位置1)开始计数的位置1至23包含至少8个核糖核苷酸;该反义链是36至66个核苷酸残基的长度,且从3'端核苷酸开始计数,在与正义链的位置1至23配对以形成双链体的位置中包含至少8个核糖核苷酸;其中至少反义链的3'末端核苷酸未与正义链配对,且至多6个连续的3'末端核苷酸未与正义链配对,从而形成具有1至6个核苷酸的3'单链突出端;其中反义链的5'末端包含10至30个为与正义链配对的连续核苷酸,从而形成具有10至30个核苷酸的单链5'突出端;其中,当将该正义链与反义链对齐以进行最大互补时,至少该正义链的5'末端核苷酸及3'末端核苷酸与反义链的核苷酸进行碱基配对,从而在该正义链与反义链之间形成实质上双链体的区域;以及,反义链在沿着反义链长度的至少19个核苷酸与靶标RNA充分互补,以在当将该双链核酸引入哺乳动物细胞内时降低靶标基因的表达;以及,其中该正义链含有至少一个位于三个连续核苷酸上的三个2'-F修饰的基序,其中这些基序的至少一者出现在裂解位点或邻近该裂解位点处。反义链含有至少一个位于裂解位点或邻近该裂解位点处的三个连续核苷酸上的三个2'-O-甲基修饰的基序。

[0356] 在一个实施方案中,RNAi剂包含正义链及反义链,其中该RNAi剂包含具有至少25个且至多29个核苷酸的长度的第一链,以及具有至多30个核苷酸的长度且具有至少一个位于从5'末端计数第11、12、13位置处的三个连续核苷酸上的三个2'-O-甲基修饰的基序的第二链;其中该第一链的3'末端及该第二链的5'末端形成钝端,且该第二链于其3'末端比该第一链长1至4个核苷酸,其中该双链体区域的长度是至少25个核苷酸,且该第二链在沿着该第二链长度的至少19个核苷酸与靶标RNA充分互补,以在当将该RNAi剂引入哺乳动物细胞内时降低靶标基因的表达,以及,其中该RNAi剂的切丁酶裂解优先得到包含该第二链的3'末端的siRNA,从而降低该哺乳动物体内的靶标基因的表达。任选的,RNAi剂进一步包含配体。

[0357] 在一个实施方案中,RNAi剂的正义链含有至少一个位于三个连续核苷酸上的三个一致修饰的基序,其中这些基序中的一个出现在正义链的裂解位点处。

[0358] 在一个实施方案中,RNAi剂的反义链还可含有至少一个位于三个连续核苷酸上的三个一致修饰的基序,其中这些基序中的一个出现在反义链的裂解位点或邻近该裂解位点处。

[0359] 对于具有17至23个核苷酸的长度的双链体区域的RNAi剂,该反义链的裂解位点典型位于从5'端计数的位置10、11、12附近。因此,这三个一致修饰的基序可出现在反义链的位置9、10、11,位置10、11、12,位置11、12、13,位置12、13、14,或位置13、14、15,从该反义链的5'端的第一个核苷酸开始计数,或在该双链体区域内从该反义链的5'端的第一个配对核苷酸开始计数。反义链的裂解位点还可根据该RNAi的双链体区域从5'端计数的长度而改变。

[0360] RNAi剂的正义链可含有至少一个位于该链的裂解位点处的三个连续核苷酸上的三个一致修饰的基序;且反义链可具有至少一个位于该链的裂解位点或邻近该裂解位点处的三个连续核苷酸上的三个一致修饰的基序。当正义链及反义链形成dsRNA双链体时,该正

义链及该反义链可经对齐，使得位于该正义链上的一个三核苷酸基序与位于该反义链上的一个三核苷酸基序具有至少一个核苷酸重叠，即，该正义链中基序的三个核苷酸的至少一者与该反义链中基序的三个核苷酸的至少一者碱基配对。另选地，至少两个核苷酸可重叠，或全部三个核苷酸可重叠。

[0361] 在一个实施方案中，RNAi剂的正义链可含有多于一个位于三个连续核苷酸上的三个一致修饰的基序。第一基序可出现在该链的裂解位点或邻近该裂解位点处，且其他基序可以是侧翼修饰。本文中，术语“侧翼修饰”是指出现在该链的另一部位的与位于相同链的裂解位点或邻近该裂解位点处的基序分隔开来的基序。侧翼修饰或与第一基序相邻或通过至少一个或多个核苷酸与第一基序分隔开来。当这些基序彼此紧邻时，则这些基序的化学性彼此截然不同；而当这些基序通过一个或多个核苷酸分隔开来时，则这些化学性可相同或不同。可存在两个或多个侧翼修饰。例如，当存在两个侧翼修饰时，每一侧翼修饰可出现在位于裂解位点或邻近该裂解位点处的第一基序的一端，或出现在该前导基序的任一侧。

[0362] 与正义链相似，RNAi剂的反义链可含有多于一个位于三个连续核苷酸上的三个一致修饰的基序，且这些基序的至少一者出现在该链的裂解位点或邻近该裂解位点处。这一反义链还可含有一个或多个侧翼修饰，该侧翼修饰的排列类似于可能存在于该正义链的侧翼修饰。

[0363] 在一个实施方案中，RNAi剂的正义链或反义链的侧翼修饰典型不包括位于该链的3’端、5’端或两端的最开始的一个或两个末端核苷酸。

[0364] 在另一实施方案中，RNAi剂的正义链或反义链的侧翼修饰典型不包括位于该双链体区域内该链的3’端、5’端或两端的最开始的一个或两个配对核苷酸。

[0365] 当RNAi剂的正义链及反义链各自含有至少一个侧翼修饰时，该侧翼修饰可落入该双链体区域的相同端，且具有一个、两个或三个核苷酸的重叠。

[0366] 当RNAi剂的正义链及反义链各自含有至少两个侧翼修饰时，该正义链与该反义链可经对齐而使得来自一链的两个修饰的每一者落入该双链体区域的一端且具有一个、两个或三个核苷酸的重叠；来自一链的两个修饰的每一者落入该双链体区域的另一端且具有一个、两个或三个核苷酸的重叠；一链的两个修饰分别落入该前导基序的两端上且在该双链体区域内具有一个、两个或三个核苷酸的重叠。

[0367] 在一个实施方案中，RNAi剂包含与靶标的错配、双链体中的错配、或其组合。错配可出现在突出端区域内或双链体区域内。基于碱基对促进解离或熔融的倾向性(例如，基于特定配对的关联或解离的自由能，自由能是基于单个配对检查配对的最简单的途径，但还可使用次近邻分析及类似分析)，可将碱基对排名。就促进解离而言：A:U优于G:C；G:U优于G:C；且I:C优于G:C(I是肌苷)。错配例如非标准配对或除标准配对之外的(如本文中他处所描述)优于标准(A:T、A:U、G:C)配对；且包括万用碱基的配对优于标准配对。

[0368] 在一个实施方案中，RNAi剂包含的位于该双链体区域内的反义链中从5’端计数最前列的第1、2、3、4或5个碱基对的至少一者，其独立地选自下列所组成的组：A:U、G:U、I:C、以及错配(例如非标准配对或除标准配对之外的的配对或包括万用碱基的配对)，以促进反义链在该双链体的5’端的解离。

[0369] 在一个实施方案中，位于该双链体区域内的反义链中从5’端计数的第1个位置处的核苷酸是选自由A、dA、dU、U及dT所组成的组。另选地，位于该双链体区域内的反义链中从

5’端计数最前列的第1、2或3个碱基对的至少一者是AU碱基对。例如，位于该双链体区域内的反义链中从5’端计数的第一个碱基对是AU碱基对。

[0370] 在另一实施方案中，位于该正义链的3’端的核苷酸是脱氧胸腺嘧啶(dT)。在另一实施方案中，位于该反义链的3’端的核苷酸是脱氧胸腺嘧啶(dT)。在一个实施方案中，在正义链或反义链的3’端存在脱氧胸腺嘧啶核苷酸的短序列，例如，两个dT核苷酸。

[0371] 在一个实施方案中，正义链序列可通过式(I)表示：

[0372] $5'n_p\text{-}N_a\text{-}(X\ X\ X)_i\text{-}N_b\text{-}Y\ Y\ Y\text{-}N_b\text{-}(Z\ Z\ Z)_j\text{-}N_a\text{-}n_q\ 3'$ (I)

[0373] 其中：

[0374] i及j各自独立为0或1；

[0375] p及q各自独立为0至6；

[0376] N_a各自独立表示包含0至25个经修饰的核苷酸的寡核苷酸序列，每一序列包含至少两个经不同修饰的核苷酸；

[0377] N_b各自独立表示包含0至10个经修饰的核苷酸的寡核苷酸序列；

[0378] n_p与n_q各自独立表示突出端核苷酸；

[0379] 其中，Nb及Y不具有相同的修饰；以及

[0380] XXX、YYY及ZZZ各自独立表示一个位于三个连续核苷酸上的三个相同修饰的基序。优选地，YYY全部为2’-F修饰的核苷酸。

[0381] 在一个实施方案中，N_a或N_b包含交替模式的修饰。

[0382] 在一个实施方案中，YYY基序出现于该正义链的裂解位点处或附近。例如，当该RNAi剂具有长度为17至23个核苷酸的双链体区域时，该YYY基序出现于该正义链的裂解位或其附近(例如，可出现于位置6、7、8；7、8、9；9、10、11；10、11、12；或11、12、13)，从5’端的第1个核苷酸开始计数；或任选的，从5’端的双链体区域内第一个配对核苷酸开始计数。

[0383] 在一个实施方案中，i是1且j是0，或i是0且j是1，或i及j两者均是1。正义链可因此通过下列式表示：

[0384] $5'n_p\text{-}N_a\text{-}YYY\text{-}N_b\text{-}ZZZ\text{-}N_a\text{-}n_q\ 3'$ (Ib)；

[0385] $5'n_p\text{-}N_a\text{-}XXX\text{-}N_b\text{-}YYY\text{-}N_a\text{-}n_q\ 3'$ (Ic)；或

[0386] $5'n_p\text{-}N_a\text{-}XXX\text{-}N_b\text{-}YYY\text{-}N_b\text{-}ZZZ\text{-}N_a\text{-}n_q\ 3'$ (Id)。

[0387] 当正义链通过式(Ib)表示时，Nb表示包含0至10、0至7、0至5、0至4、0至2或0个经修饰的核苷酸的寡核苷酸序列。

[0388] 各N_a独立可表示包含2至20、2至15或2至10个经修饰的核苷酸的寡核苷酸序列。

[0389] 当正义链通过式(Ic)表示时，N_b表示包含0至10、0至7、0至10、0至7、0至5、0至4、0至2或0个经修饰的核苷酸的寡核苷酸序列。N_a可各自独立表示包含2至20、2至15或2至10个经修饰的核苷酸的寡核苷酸序列。

[0390] 当正义链通过式(Id)表示时，N_b各自独立表示包含0至10、0至7、0至5、0至4、0至2或0个经修饰的核苷酸的寡核苷酸序列。优选地，N_b是0、1、2、3、4、5或6。N_a可各自独立表示包含2至20、2至15或2至10个经修饰的核苷酸的寡核苷酸序列。

[0391] X、Y及Z各者可以是彼此相同或不同。

[0392] 在其他实施方案中，i是0且j是0，且正义链可通过下式表示：

[0393] $5'n_p\text{-}N_a\text{-}YYY\text{-}N_a\text{-}n_q\ 3'$ (Ia)。

[0394] 当正义链通过式(Ia)表示时，N_a各自独立表示包含2至20、2至15或2-10个经修饰的核苷酸的寡核苷酸序列。

[0395] 在一个实施方案中，RNAi的反义链序列可通过式(II)表示：

[0396] $5'n_{q'}-N_a'-(Z'Z'Z')_k-N_b'-Y'Y'Y'-N_b'-(X'X'X')_l-N'_a-n_p'3'$ (II)

[0397] 其中：

[0398] k及l各自独立为0或1；

[0399] p'及q'各自独立为0至6；

[0400] N_a'各自独立表示包含0至25个经修饰的核苷酸的寡核苷酸序列，每一序列包含至少两个经不同修饰的核苷酸；

[0401] N_b'各自独立表示包含0至10个经修饰的核苷酸的寡核苷酸序列；

[0402] n_p'与n_q'各自独立表示突出端核苷酸；

[0403] 其中，N_b'及Y'不具有相同的修饰；

[0404] 以及，X'X'X'、Y'Y'Y'及Z'Z'Z'各自独立表示一个位于三个连续核苷酸上的三个相同修饰的基序。

[0405] 在一个实施方案中，N_a'或N_b'包含交替模式的修饰。

[0406] 在一个实施方案中，Y'Y'Y'基序位于反义链的裂解位点处或临近反义链的裂解位点。例如，当该RNAi剂是具有长度为17至23个核苷酸的双链体区域时，该Y'Y'Y'基序是出现于该反义链的位置9、10、11；10、11、12；11、12、13；12、13、14；或13、14、15，从5'端的第1个核苷酸开始计数；或任选的，从5'端的双链体区域内第一个配对核苷酸开始计数。优选地，Y'Y'Y'基序出现于位置11、12、13处。

[0407] 在一个实施方案中，Y'Y'Y'基序全部为2'-OMe修饰的核苷酸。

[0408] 在一个实施方案中，k是1且l是0，或k是0且l是1，或k及l两者均是1。

[0409] 反义链可因此通过下列式表示：

[0410] $5'n_{q'}-N_a'-Z'Z'Z'-N_b'-Y'Y'Y'-N_a'-n_{p'}3'$ (IIb)；

[0411] $5'n_{q'}-N_a'-Y'Y'Y'-N_b'-X'X'X'-n_{p'}3'$(IIc)；或

[0412] $5'n_{q'}-N_a'-Z'Z'Z'-N_b'-Y'Y'Y'-N_b'-X'X'X'-N_a'-n_{p'}3'$ (IId)。

[0413] 当反义链通过式(IIb)表示时，N_b'表示包含0至10、0至7、0至10、0至7、0至5、0至4、0至2或0个经修饰的核苷酸的寡核苷酸序列。各N_a'独立表示包含2至20、2至15或2至10个经修饰的核苷酸的寡核苷酸序列。

[0414] 当反义链表示为式(IIc)时，N_b'表示包含0至10、0至7、0至10、0至7、0至5、0至4、0至2或0个经修饰的核苷酸的寡核苷酸序列。各N_a'独立表示包含2至20、2至15或2至10个经修饰的核苷酸的寡核苷酸序列。

[0415] 当反义链表示为式(IId)时，各N_b'独立表示包含0至10、0至7、0至10、0至7、0至5、0至4、0至2或0个经修饰的核苷酸的寡核苷酸序列。各N_a'独立表示包含2至20、2至15或2至10个经修饰的核苷酸的寡核苷酸序列。优选地，N_b是0、1、2、3、4、5或6。

[0416] 在其他实施方案中，k是0且l是0，且反义链可通过下式表示：

[0417] $5'n_{p'}-N_{a'}-Y'Y'Y'-N_{a'}-n_{q'}3'$ (Ia)。

[0418] 当反义链表示为式(Ia)时，N_a'各自独立表示包含2至20、2至15或2至10个经修饰的核苷酸的寡核苷酸序列。

[0419] X'、Y'及Z'各者可以是彼此相同或不同。

[0420] 正义链及反义链的每一核苷酸独立经LNA、HNA、CeNA、2'-己氧基乙基、2'-O-甲基、2'-O-烯丙基、2'-C-烯丙基、2'-羟基或2'-氟修饰。例如,正义链及反义链的每一核苷酸独立经2'-O-甲基或2'-氟修饰。特别地,各X、Y、Z、X'、Y'及Z'可表示2'-O-甲基修饰或2'-F修饰。

[0421] 在一个实施方案中,当双链体区域是21nt时,该RNAi剂的正义链可含有出现在该链的9、10及11位置的YYY基序,从5'末端的第一个核苷酸开始计数,或任选的,在双链体区域内从5'末端的第一个配对核苷酸开始计数;以及,Y是表示2'-F修饰。正义链可额外含有XXX基序或ZZZ基序作为位于双链体区域的相反末端的侧翼修饰;以及,XXX与ZZZ各自独立表示2'-OMe修饰或2'-F修饰。

[0422] 在一个实施方案中,反义链可含有出现在该链的11、12及13位置的Y′Y′Y′基序,从5'末端的第一个核苷酸开始计数,或任选的,在双链体区域内从5'端的第一个配对核苷酸开始计数;以及,Y′是表示2'-O-甲基修饰。反义链可额外含有X'X'X'基序或Z'Z'Z'基序作为位于双链体区域的相反末端的侧翼修饰;以及,X'X'X'与Z'Z'Z'各自独立表示2'-OMe修饰或2'-F修饰。

[0423] 由上述式(Ia)、(Ib)、(Ic)及(Id)中任一者表示的正义链分别与由式(IIa)、(IIb)、(IIc)及(IId)中任一者表示的反义链形成双链体。

[0424] 据此,用于本公开的方法中的RNAi剂可包含正义链及反义链,每一条链具有14至30个核苷酸,该RNAi双链体由式(III)表示:

[0425] 正义:5'n_p-N_a-(X X X)$_i$-N_b-Y Y Y-N_b-(Z Z Z)$_j$-N_a-n_q 3'

[0426] 反义:3'n_p'-N_a'-(X'X'X')$_k$-N_b'-Y'Y'Y'-N_b'-(Z'Z'Z')$_l$-N_a'-n_q'5'

[0427] (III)

[0428] 其中:

[0429] i、j、k及l各自独立为0或1;

[0430] p、p'、q及q'各自独立为0至6;

[0431] N_a及N_a'各自独立表示包含0至25个经修饰的核苷酸的寡核苷酸序列,每一序列包含至少两个经不同修饰的核苷酸;

[0432] N_b与N_b′各自独立表示包含0至10个经修饰的核苷酸的寡核苷酸序列。

[0433] 其中

[0434] n_p'、n_p、n_q'及n_q各自可存在或不存在,且各自独立表示突出端核苷酸;以及

[0435] XXX、YYY、ZZZ、X'X'X'、Y'Y'Y'及Z'Z'Z'各自独立表示一个位于三个连续核苷酸上的三个相同修饰的基序。

[0436] 在一个实施方案中,i是0且j是0;或i是1且j是0;或i是0且j是1;或i及j两者均是0;或i及j两者均是1。在另一实施方案中,k是0且l是0;或k是1且l是0;或k是0且l是1;或k及l两者均是0;或k及l两者均是1。

[0437] 形成RNAi双链体的正义链与反义链的示例性组合包括下述各式:

[0438] 5'n_p-N_a-Y Y Y-N_a-n_q 3'

[0439] 3'n_p'-N_a'-Y′Y′Y′-N_a'n_q'5'

[0440] (IIIa)

[0441] 5'n_p-N_a-Y Y Y-N_b-Z Z Z-N_a-n_q 3'

[0442] 3'n_p'-N_a'-Y'Y'Y'-N_b'-Z'Z'Z'-N_a'n_q'5'

[0443] (IIIb)

[0444] 5'n_p-N_a-X X X-N_b-Y Y Y-N_a-n_q 3'

[0445] 3'n_p'-N_a'-X'X'X'-N_b'-Y'Y'Y'-N_a'-n_q'5'

[0446] (IIIc)

[0447] 5'n_p-N_a-X X X-N_b-Y Y Y-N_b-Z Z Z-N_a-n_q 3'

[0448] 3'n_p'-N_a'-X'X'X'-N_b'-Y'Y'Y'-N_b'-Z'Z'Z'-N_a-n_q'5'

[0449] (IIId)

[0450] 当RNAi剂是通过式(IIIa)表示时,各N_a独立表示包含2至20、2至15或2至10个经修饰的核苷酸的寡核苷酸序列。

[0451] 当RNAi剂是通过式(IIIb)表示时,各N_b独立表示包含1至10、1至7、1至5或1至4个经修饰的核苷酸的寡核苷酸序列。各N_a独立表示包含2至20、2至15或2至10个经修饰的核苷酸的寡核苷酸序列。

[0452] 当RNAi剂表示为式(IIIc)时,各N_b、N_b'独立表示包含0至10、0至7、0至10、0至7、0至5、0至4、0至2或0个经修饰的核苷酸的寡核苷酸序列。各N_a独立表示包含2至20、2至15或2至10个经修饰的核苷酸的寡核苷酸序列。

[0453] 当RNAi剂表示为式(IIId)时,各N_b、N_b'独立表示包含0至10、0至7、0至10、0至7、0至5、0至4、0至2或0个经修饰的核苷酸的寡核苷酸序列。各N_a、N_a'各独立表示包含2至20、2至15或2至10个经修饰的核苷酸的寡核苷酸序列。N_a、N_a'、N_b及N_b'各者独立包含交替模式的修饰。

[0454] 在一个实施方案中,当RNAi剂由式(IIId)表示时,N_a修饰是2'-O-甲基修饰或2'氟修饰。在另一实施方案中,当RNAi剂由式(IIId)表示时,N_a修饰是2'-O-甲基修饰或2'-氟修饰,且n_p'>0,以及,至少一个n_p'经由硫代磷酸酯键而键联至相邻核苷酸。在又一实施方案中,当RNAi剂由式(IIId)表示时,N_a修饰是2'-O-甲基修饰或2'氟修饰,n_p'>0,且至少一个n_p'经由硫代磷酸酯键而键联至相邻核苷酸,以及,正义链通过二价或三价分支的接头(下文描述)缀合至一个多个C16(或相关)部分。在另一实施方案中,当RNAi剂由式(IIId)表示时,N_a修饰是2'-O-甲基修饰或2'氟修饰,n_p'>0,且至少一个n_p'经由硫代磷酸酯键而键联至相邻核苷酸,正义链包含至少一个硫代磷酸酯键,以及,正义链任选地通过二价或三价分支的接头(下文描述)附接而缀合至一个多个亲脂性(例如C16(或相关))部分。

[0455] 在一个实施方案中,当RNAi剂由式(IIIa)表示时,N_a修饰是2'-O-甲基修饰或2'氟修饰,n_p'>0,且至少一个n_p'经由硫代磷酸酯键而键联至相邻核苷酸,正义链包含至少一个硫代磷酸酯键,以及,正义链通过二价或三价分支的接头(下文描述)附接而缀合至一个多个亲脂性例(如C16(或相关))部分。

[0456] 在一个实施方案中,RNAi剂是含有至少两个由式(III)、(IIIa)、(IIIb)、(IIIc)及(IIId)表示的双链体的多聚物,其中这些双链体通过接头连接。接头可以是可裂解的或不可裂解的。任选的,多聚物进一步包含配体。双链体可各自靶向相同基因或两个不同基因;或双链体可各自靶向相同基因的两个不同靶标位点。

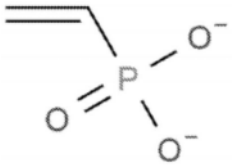
[0457] 在一个实施方案中,RNAi剂含有3、4、5、6或更多个由式(III)、(IIIa)、(IIIb)、

(IIIc)及(IIId)表示的双链体的多聚物,其中这些双链体通过接头连接。接头可以是可裂解者或不可裂解者。任选的,多聚物进一步包含配体。双链体可各自靶向相同基因或两个不同基因;或双链体可各自靶向相同基因的两个不同靶标位点。

[0458] 在一个实施方案中,两个由式(III)、(IIIa)、(IIIb)、(IIIc)及(IIId)表示的RNAi剂在5'端及一个或两个3'端彼此键联,且任选的缀合至配体。这些剂可各自靶向相同基因或两个不同基因;或这些剂可各自靶向相同基因的两个不同靶标位点。

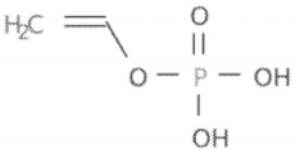
[0459] 多个出版物描述可用于本公开的方法中的多聚RNAi剂。此类出版物包括WO2007/091269、WO2010/141511、WO2007/117686、WO2009/014887、WO2011/031520及US 7858769,其各自的全部内容是通过援引而并入本文。

[0460] 在某些实施方案中,本公开的组合物及方法包括本文所描述的RNAi剂的膦酸乙烯酯(VP)修饰。在示例性实施方案中,本公开的膦酸乙烯酯具有下述结构:

[0461] 

[0462] 本公开的膦酸乙烯酯可附接至本公开的dsRNA的反义链或正义链。在某些优选实施方案中,本公开的膦酸乙烯酯附接至dsRNA的反义链,任选的附接在dsRNA的反义链的5'端。

[0463] 膦酸乙烯酯修饰也预期用于本公开的组合物及方法中。示例性膦酸乙烯酯结构是:

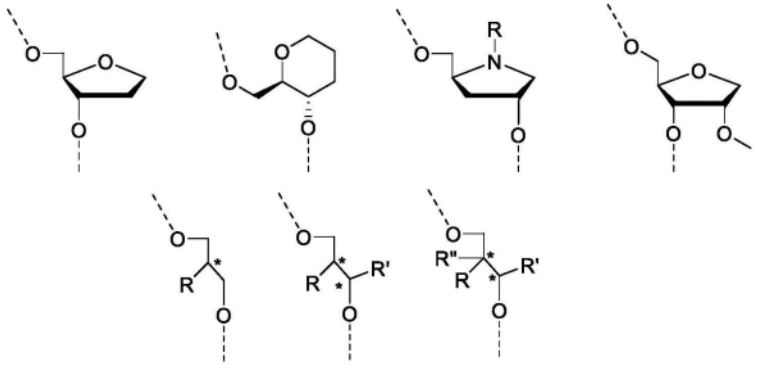
[0464] 

[0465] i.热去稳定化修饰

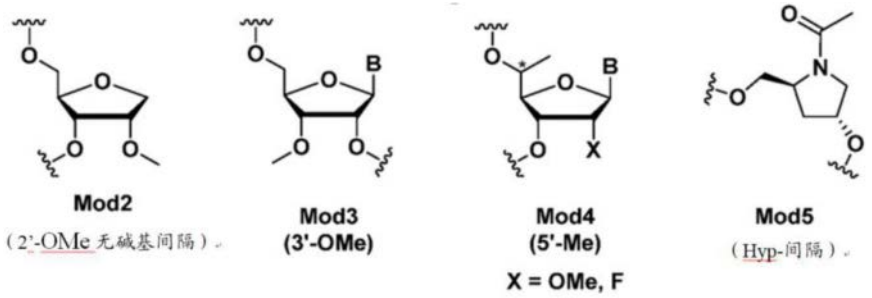
[0466] 在某些实施方案中,通过将热去稳定化修饰并入反义链的种子区域(即,反义链的5'端的位置2至9)以降低或抑制脱靶基因沉默,从而优化dsRNA分子以用于RNA干扰。已经发现,具有包含位于从反义链5'末端计数最先9个核苷酸内的双链体的至少一个热去稳定化修饰的反义链的dsRNA具有降低的脱靶基因沉默活性。据此,在一些实施方案中,反义链包含位于从反义链5'末端计数最先9个核苷酸内的双链体的至少一个(例如,一个、两个、三个、四个、五个或更多个)热去稳定化修饰。在一些实施方案中,双链体的一个或多个热去稳定化修饰是位于从反义链5'末端计数的位置2至9处,或优选位置4至8处。在又一些实施方案中,双链体的热去稳定化修饰是位于从反义链5'端计数的位置6、7或8处。在再一些实施方案中,双链体的热去稳定化修饰是位于从反义链5'端计数的位置7处。术语"热去稳定化修饰"包括将导致dsRNA具有较低总体熔融温度(Tm)(优选地,Tm比不具有此类修饰的dsRNA的Tm低一度、两度、三度或四度)的修饰。在一些实施方案中,双链体的热去稳定化修饰是位于从反义链5'末端计数的位置2、3、4、5或9处。

[0467] 热去稳定化修饰可包括但不限于,无碱基的修饰;与相反链中相反核苷酸的错配;以及糖修饰诸如2'-脱氧修饰或非环状核苷酸,例如未锁核酸(UNA)或二醇核酸(GNA)。

[0468] 示例性的无碱基的修饰包括但不限于下列:

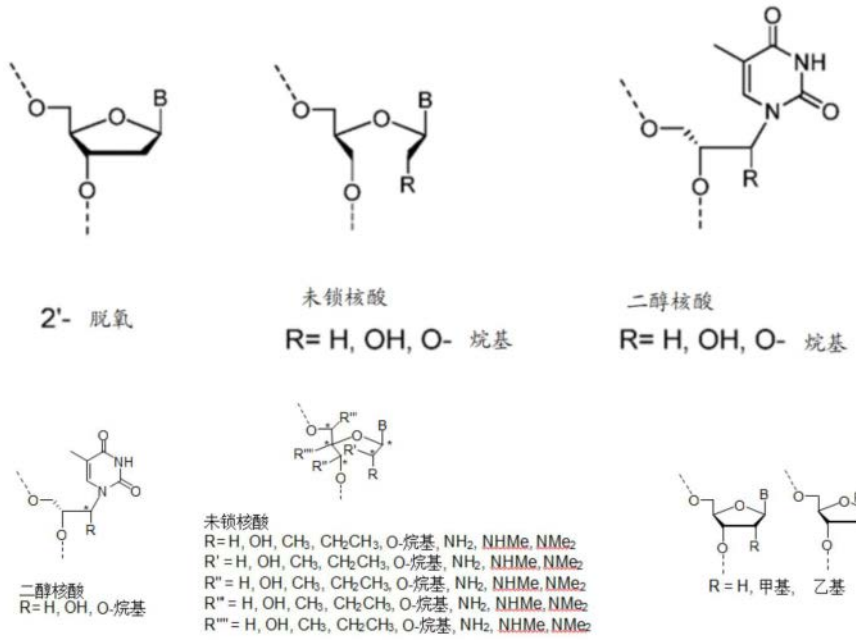
[0469] 

[0470] 其中,R=H、Me、Et或OMe;R'=H、Me、Et或OMe;R"=H、Me、Et或OMe

[0471] 

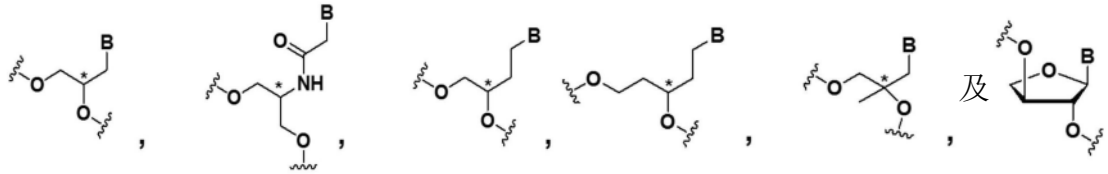
[0472] 其中,B是经修饰或未修饰的核碱基。

[0473] 示例性的糖修饰包括但不限于下列:

[0474] 

[0475] 其中,B是经修饰或未修饰的核碱基。

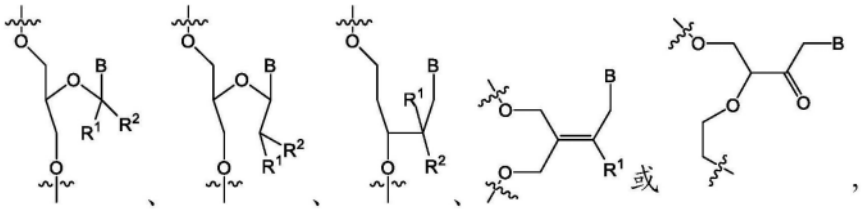
[0476] 在一些实施方案中,双链体的热去稳定化修饰是选自由下列所组成的组:

[0477] 

[0478] 其中,B是经修饰或未修饰的核碱基,并且每个结构上的星号表示R、S或外消旋。

[0479] 术语"非环状核苷酸"是指具有非环状核糖的任何核苷酸,例如,其中核糖碳之间

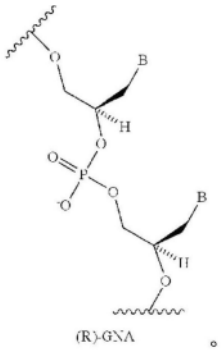
的任何键(例如,C1'-C2'、C2'-C3'、C3'-C4'、C4'-O4'或C1'-O4')是不存在或核糖碳或氧的至少一者(例如,C1'、C2'、C3'、C4'或O4')独立地或组合地不存在于该核苷酸中。在一些实施方案中,非环状核苷酸是



其中,B是经修饰或未修饰的核碱基,R^1及R^2独立为H、卤素、OR_3或烷基;以及,R_3是H、烷基、环烷基、芳基、芳烷基、杂芳基或糖)。术语"UNA"是指未锁定的非环状核酸,其中该糖的任何键已经移除,形成未锁定的"糖"残基。在一示例中,UNA亦涵盖其C1'-C4'键(即,位于C1'碳与C4'碳间的碳-氧-碳共价键)被移除的单体。在另一示例中,糖的C2'-C3'键(即,C2'碳与C3'碳之间的碳-碳共价键)已经移除(参见,Mikhailov et.al.,Tetrahedron Letters,26(17):2059(1985);以及Fluiter et al.,Mol.Biosyst.,10:1039(2009),其均通过援引而整体并入本文)。非环状衍生物提供更大的骨架柔性而不影响Watson-Crick配对。非环状核苷酸可经由2'-5'或3'-5'键而键联。

[0480] 术语"GNA"是指二醇核酸,其是类似于DNA或RNA但其"骨架"组成(由通过磷酸二酯键键的重复甘油单元构成)上与DNA或RNA有所区别的聚合物:

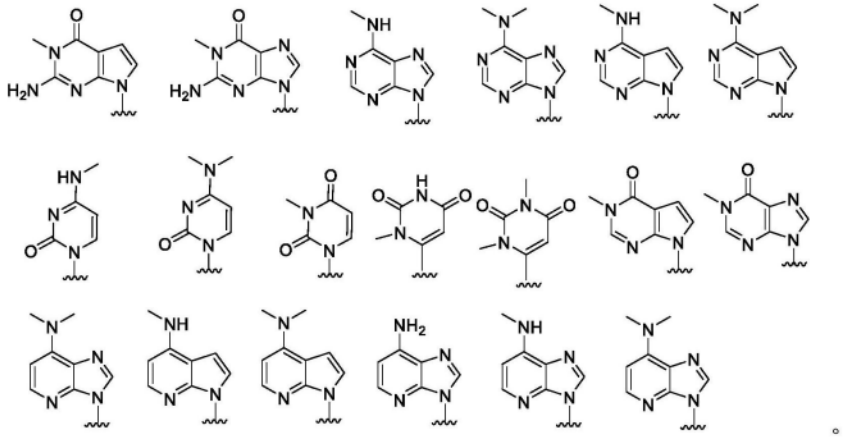
[0481]



。

[0482] 双链体的热去稳定化修饰可以是热去稳定化核苷酸与dsRNA双链体内相反链的相反核苷酸之间的错配(即,非互补性碱基对)。示例性错配碱基对包括G:G、G:A、G:U、G:T、A:A、A:C、C:C、C:U、C:T、U:U、T:T、U:T、或其组合。本领域中已知的其他错配碱基配对也适用于本发明。错配可出现于天然出现的核苷酸之间或经修饰的核苷酸之间,即,错配碱基配对可出现于来自核苷酸中核糖的独立修饰的相应核苷酸的核碱基之间。在某些实施方案中,dsRNA分子含有至少一个错配对中的核碱基,其是2'-脱氧核碱基;例如,该2'-脱氧核碱基位于正义链中。

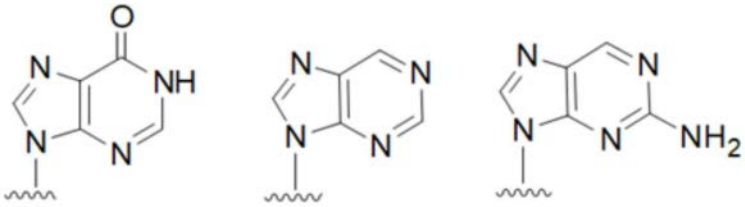
[0483] 在一些实施方案中,反义链的种子区域内的双链体的热去稳定化修饰包括其与靶标mRNA上的互补碱基的W-C H键键合受损的核苷酸,诸如:

[0484] 

[0485] 无碱基的核苷酸、非环状核苷酸修饰(包括UNA及GNA)及错配修饰的更多示例已经详细描述于WO 2011/133876中,该专利通过援引而以其整体并入本文。

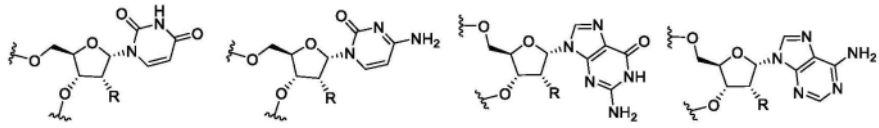
[0486] 热去稳定化修饰还可包括通用碱基及磷酸酯修饰,该通用碱基与相反碱基形成氢键的能力被降低或废除。

[0487] 在一些实施方案中,双链体的热去稳定化修饰包括具有非标准碱基的核苷酸,诸如但不限于,其与相反链中的碱基形成氢键的能力受损或被完全废除的核苷酸修饰。这些核碱基修饰已经评估其对于dsRNA双链体的中心区域的去稳定化,如WO 2010/0011895中所描述,该专利通过援引而以其整体并入本文。示例性核碱基修饰是:

[0488] 

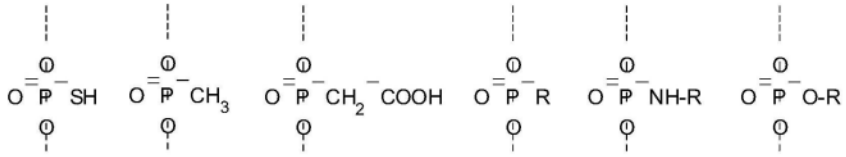
[0489] 

[0490] 在一些实施方案中,反义链的种子区域内的双链体的热去稳定化修饰包括与靶标mRNA上的互补的一个或多个α-核苷酸,诸如:

[0491] 

[0492] 其中,R是H、OH、OCH_3、F、NH_2、NHMe、NMe_2或O-烷基。

[0493] 已知相对于天然磷酸二酯键降低dsRNA双链体的热稳定性的示例性磷酸酯修饰是:

[0494] 

R = 烷基

[0495] R基团的烷基可以是C_1-C_6烷基。用于R基团的具体烷基包括但不限于甲基、乙基、丙基、异丙基、丁基、戊基及己基。

[0496] 本领域技术人员应认识到,鉴于核碱基的功能角色是定义本公开的RNAi剂的特异性,而核碱基修饰可以如本文所描述的各种方式执行(例如以将去稳定化修饰引入本公开的RNAi剂,例如用于相对于脱靶效应而增强上靶效应的目的),可用并且通常于本公开的RNAi剂的修饰的范围倾向于比非核碱基修饰(例如对多核糖核苷酸的糖基团或磷酸酯骨架的修饰)大得多。此类修饰更详细地描述于本公开的其他章节中,并且明确为本公开的RNAi剂所考虑,不论是具备天然核碱基的还是具备上述或本文中他处所描述的修饰核碱基的。

[0497] 除了包含热去稳定化修饰的反义链之外,dsRNA还可包含一个或多个稳定化修饰。例如,dsRNA可包含至少两个(例如,两个、三个、四个、五个、六个、七个、八个、九个、十个或更多个)稳定化修饰。非限制性地,稳定化修饰全部可存在于一条链中。在一些实施方案中,正义链及反义链两者都包含至少两个稳定化修饰。稳定化修饰可出现于正义链或反义链的任何核苷酸上。例如,稳定化修饰可出现在正义链和/或反义链的每一个核苷酸;每一稳定化修饰可以交替模式出现在正义链或反义链上;或正义链或反义链包含交替模式的两种稳定化修饰。正义链的交替模式的稳定化修饰可与反义链相同或不同,并且正义链交替模式的稳定化修饰可具有相对于反义链的交替模式的稳定化修饰的变动。

[0498] 在一些实施方案中,反义链包含至少两个(例如,两个、三个、四个、五个、六个、七个、八个、九个、十个或更多个)稳定化修饰。非限制性地,反义链中的稳定化修饰可存在于任何点处。在一些实施方案中,反义链包含位于自5’端计数的位置2、6、8、9、14及16处的稳定化修饰。在一些其他实施方案中,反义链包含位于自5’端计数的位置2、6、14及16处的稳定化修饰。在再一些其他实施方案中,反义链包含位于自5’末端计数的位置2、14及16处的稳定化修饰。

[0499] 在一些实施方案中,反义链包含与去稳定化修饰相邻的至少一个稳定化修饰。例如,稳定化修饰可以是位于去稳定化修饰的5’端或3’端的核苷酸,即,位于自该去稳定化修饰位置起-1或+1位置处。在一些实施方案中,反义链包含稳定化修饰,该稳定化修饰位于去稳定化修饰的5’端及3’端中的每一处,即,位于自该去稳定化修饰位置起-1及+1位置处。

[0500] 在一些实施方案中,反义链包含至少两个稳定化修饰,该稳定化修饰位于去稳定化修饰的3’端,即,位于自该去稳定化修饰位置起+1及+2位置处。

[0501] 在一些实施方案中,正义链包含至少两个(例如,两个、三个、四个、五个、六个、七个、八个、九个、十个或更多个)稳定化修饰。非限制性地,正义链中的稳定化修饰可存在于任何点处。在一些实施方案中,正义链包含位于自5’端计数的位置7、10及11处的稳定化修饰。在一些其他实施方案中,正义链包含位于自5’端计数的位置7、9、10及11处的稳定化修饰。在一些实施方案中,正义链包含稳定化修饰,该稳定化修饰位于与反义链的自该反义链5’端计数的位置11、12及15相反或互补的位置处。在一些其他实施方案中,正义链包含稳定

化修饰，该稳定化修饰位于与反义链的自该反义链5’端计数的位置11、12、13及15相反或互补的位置处。在一些实施方案中，正义链包含两个、三个或四个稳定化修饰的嵌段。

[0502] 在一些实施方案中，正义链不包含位于与反义链中该双链体的热去稳定化修饰相反或互补的位置处的稳定化修饰。

[0503] 示例性的热稳定化修饰包括但不限于，2’-氟修饰。其他热稳定化修饰包括但不限于，LNA。

[0504] 在一些实施方案中，本公开的dsRNA包含至少四个(例如，四个、五个、六个、七个、八个、九个、十个或更多个)2’-氟核苷酸。非限制性地，2’-氟核苷酸全部可存在于一条链中。在一些实施方案中，正义链及反义链两者都包含至少两个2’-氟核苷酸。2’-氟修饰可出现于正义链或反义链的任何核苷酸上。例如，2’-氟修饰可出现在正义链和/或反义链的每一个核苷酸上；每一2’-氟修饰可以交替模式出现在正义链或反义链上；或正义链或反义链包含交替模式的两种2’-氟修饰。正义链的交替模式的2’-氟修饰可与反义链相同或不同，其正义链的交替模式的2’-氟修饰可具有相对于反义链的交替模式的2’-氟修饰的变动。

[0505] 在一些实施方案中，反义链包含至少两个(例如，两个、三个、四个、五个、六个、七个、八个、九个、十个或更多个)2’-氟核苷酸。非限制性地，反义链中的2’-氟修饰可存在于任何点处。在一些实施方案中，反义链包含位于自5’端计数的位置2、6、8、9、14及16处的2’-氟核苷酸。在一些其他实施方案中，反义链包含位于自5’端计数的位置2、6、14及16处的2’-氟核苷酸。在再一些其他实施方案中，反义链包含位于自5’末端计数的位置2、14及16处的2’-氟核苷酸。

[0506] 在一些实施方案中，反义链包含与去稳定化修饰相邻的至少一个2’-氟核苷酸。例如，2’-氟核苷酸可以是位于去稳定化修饰的5’端或3’端的核苷酸，即，位于自该去稳定化修饰位置起-1或+1位置处。在一些实施方案中，反义链包含2’-氟核苷酸，该2’-氟核苷酸位于去稳定化修饰的5’端及3’端中的每一处，即，位于自该去稳定化修饰位置起-1及+1位置处。

[0507] 在一些实施方案中，反义链包含至少两个2’-氟核苷酸，该2’-氟核苷酸位于去稳定化修饰的3’端，即，位于自该去稳定化修饰位置起+1及+2位置处。

[0508] 在一些实施方案中，正义链包含至少两个(例如，两个、三个、四个、五个、六个、七个、八个、九个、十个或更多个)2’-氟核苷酸。非限制性地，正义链中的2’-氟修饰可存在于任何点处。在一些实施方案中，正义链包含位于自5’端计数的位置7、10及11处的2’-氟核苷酸。在一些其他实施方案中，正义链包含位于自5’端计数的位置7、9、10及11处的2’-氟核苷酸。在一些实施方案中，正义链包含2’-氟核苷酸，该2’-氟核苷酸位于与反义链的自该反义链5’端计数的位置11、12及15相反或互补的位置处。在一些其他实施方案中，正义链包含2’-氟核苷酸，该2’-氟核苷酸位于与反义链的自该反义链5’端计数的位置11、12、13及15相反或互补的位置处。在一些实施方案中，正义链包含两个、三个或四个2’-氟核苷酸的嵌段。

[0509] 在一些实施方案中，正义链不包含位于与反义链中该双链体的热去稳定化修饰相反或互补的位置处的2’-氟核苷酸。

[0510] 在一些实施方案中，本公开的dsRNA分子包含21个核苷酸的正义链及23个核苷酸的反义链，其中该反义链含有至少一个热去稳定化核苷酸，其中该至少一个热去稳定化核苷酸出现于反义链的种子区域中(即，位于反义链5’端的位置2至9处)，其中dsRNA的一端是

钝端，而另一端包含2nt的突出端，即，其中dsRNA任选的进一步具有下述特征的至少一个(例如，一个、两个、三个、四个、五个、六个或全部七个)：(i)反义链包含2、3、4、5或6个2'-氟修饰；(ii)反义链包含1、2、3、4或5个硫代磷酸酯核苷酸间键；(iii)正义链与配体缀合；(iv)正义链包含2、3、4或5个2'-氟修饰；(v)正义链包含1、2、3、4或5个硫代磷酸酯核苷酸间键；(vi)dsRNA包含至少四个2'-氟修饰；以及(vii)dsRNA包含位于反义链5'端的钝端。优选地，该2nt的突出端位于反义链的3'端。

[0511] 在一些实施方案中，本公开的dsRNA分子包含正义链及反义链，其中：该正义链是25至30个核苷酸残基的长度，其中该链的从5'末端核苷酸(位置1)开始计数的位置1至23包含至少8个核糖核苷酸；该反义链是36至66个核苷酸残基的长度，且从3'末端核苷酸开始计数，在与正义链的位置1至23配对以形成双链体的位置中包含至少8个核糖核苷酸；其中至少反义链的3'末端核苷酸未与正义链配对，且至多6个连续的3'末端核苷酸未与正义链配对，从而形成具有1至6个核苷酸的3'单链突出端；其中反义链的5'末端包含10至30个未与正义链配对的连续核苷酸，从而形成具有10至30个核苷酸的单链5'突出端；其中，当将该正义链与反义链对齐以进行最大互补时，至少该正义链的5'末端核苷酸及3'末端核苷酸与反义链的核苷酸进行碱基配对，从而在该正义链与反义链之间形成实质上双链体的区域；以及，反义链在沿着反义链长度的至少19个核苷酸上与靶标RNA充分互补，以在当将所述双链核酸引入哺乳动物细胞内时降低靶标基因的表达；以及，其中该反义链含有至少一个热去稳定化核苷酸，其中至少一个热去稳定化核苷酸位于反义链的种子区域内(即，位于反义链5'端的位置2至9处)。例如，热去稳定化核苷酸出现于与正义链的5'端的位置14至17相反或互补的位置之间，并且其中dsRNA任选的进一步具有下述特征的至少一个(例如，一个、两个、三个、四个、五个、六个或全部七个)：(i)反义链包含2、3、4、5或6个2'-氟修饰；(ii)反义链包含1、2、3、4或5个硫代磷酸酯核苷酸间键；(iii)正义链与配体缀合；(iv)正义链包含2、3、4或5个2'-氟修饰；(v)正义链包含1、2、3、4或5个硫代磷酸酯核苷酸间键；(vi)dsRNA包含至少四个2'-氟修饰；以及(vii)dsRNA包含长度为12至30个核苷酸对的双链体区域。

[0512] 在一些实施方案中，本公开的dsRNA分子包含正义链及反义链，其中所述dsRNA分子包含长度为至少25个核苷酸并且最多为29个核苷酸的正义链以及长度至多为30个核苷酸的反义链，且正义链包含位于自5'端起位置11处的对于酶降解敏感的修饰核苷酸，其中该正义链的3'端及该反义链的5'端形成钝端，并且所述反义链于其3'端比正义链长1至4个核苷酸，其中双链体区域的长度为至少25个核苷酸，并且当将所述dsRNA分子引入哺乳动物细胞内时，该反义链沿着该反义链的至少19nt与靶标mRNA充分互补以降低靶标基因表达，并且其中该dsRNA的切丁酶裂解优先导致包含该反义链的该3'端的siRNA，从而降低哺乳动物体内靶标基因的表达，其中该反义链含有至少一个热去稳定化核苷酸，其中该至少一个热去稳定化核苷酸位于反义链的种子区域中(即，位于反义链5'端的位置2至9处)，并且其中dsRNA任选的进一步具有下述特征的至少一个(例如，一个、两个、三个、四个、五个、六个或全部七个)：(i)反义链包含2、3、4、5或6个2'-氟修饰；(ii)反义链包含1、2、3、4或5个硫代磷酸酯核苷酸间键；(iii)正义链与配体缀合；(iv)正义链包含2、3、4或5个2'-氟修饰；(v)正义链包含1、2、3、4或5个硫代磷酸酯核苷酸间键；(vi)dsRNA包含至少四个2'-氟修饰；以及(vii)dsRNA具有长度为12至29个核苷酸对的双链体区域。

[0513] 在一些实施方案中，dsRNA分子的正义链及反义链中的每个核苷酸可经修饰。每一

核苷酸可通过相同或不同的修饰而经修饰，这些修饰可包括非连接性磷酸酯氧中的一个或两者、或连接性磷酸酯氧中的一个或多者的一个或多个变更；核糖的构建的变更，如核糖的2'羟基的变更；以"去磷"接头进行的磷酸酯部分的整体置换；天然出现的碱基的修饰或置换；以及核糖-磷酸酯骨架的置换或修饰。

[0514] 由于核酸是子单元的聚合物，这些修饰的多数出现在核酸内重复的位置，如，碱基、磷酸酯部分、或磷酸酯部分的非连接性O的修饰。在一些情形中，修饰将出现在该核酸的所有受试位置，但在多数情形中并非如此。例如，修饰可仅出现在3'或5'末端位置，可仅出现在末端区域，例如，出现在末端核苷酸的位置或出现在一条链的最末2、3、4、5或10个核苷酸内。修饰可出现在双链区域内、单链区域内、或两者内。修饰可仅出现在RNA的双链区域内，或仅出现在RNA的单链区域内。例如，位于非连接性O位置的硫代磷酸酯修饰可仅出现在一端或两端；可仅出现在末端区域，例如出现在末端核苷酸的位置或出现在一条链的最末2、3、4、5或10个核苷酸内；或可出现在双链区域及单链区域内，尤其是在末端。一个或多个5'端可经磷酸化。

[0515] 下述是可能的，例如，提升稳定性，在突出端中包括特定的碱基，或在单链突出如5'突出端或3'突出端或两者中包括经修饰的核苷酸或核苷酸替代品。例如，可能希望的是在突出中包括嘌呤核苷酸。在一些实施方案中，3'或5'突出端中的全部或一些碱基可经修饰，如具有本文所述的修饰。修饰可包括，例如，使用在核糖的2'位置具有本领域中已知的修饰，如使用脱氧核糖核苷酸，使用2'-脱氧-2'-氟(2'-F)或2'-O-甲基修饰替代核碱基的核糖，以及使用磷酸酯基团中的修饰如硫代磷酸酯修饰。突出端无需与靶标序列同源。

[0516] 在一些实施方案中，该正义链及反义链的每一残基独立经LNA、HNA、CeNA、2'-甲氧基乙基、2'-O-甲基、2'-O-烯丙基、2'-C-烯丙基、2'-脱氧或2'-氟修饰。该链可含有多于一个的修饰。在一些实施方案中，该正义链及反义链的每一残基独立经2'-O-甲基或2'-氟修饰。应理解，这些修饰是除存在于反义链中的双链体的至少一个热去稳定化修饰之外的。

[0517] 至少两个不同的修饰典型存在于正义链及反义链上。这两个修饰可以是2'-脱氧、2'-O-甲基或2'-氟修饰、非环状核苷酸等。在一些实施方案中，正义链及反义链各自包含选自2'-O-甲基或2'-脱氧的两个不同的修饰核苷酸。在一些实施方案中，正义链及反义链的每个残基独立地经下述修饰：2'-O-甲基核苷酸、2'-脱氧核苷酸、2'-脱氧-2'-氟核苷酸、2'-O-N-甲基乙酰氨基(2'-O-NMA)核苷酸、2'-O-二甲基氨基乙氧基乙基(2'-O-DMAEOE)核苷酸、2'-O-氨基丙基(2'-O-AP)核苷酸或2'-ara-F核苷酸。再次地，应理解，这些修饰是除存在于反义链中的双链体的至少一个热去稳定化修饰之外的。

[0518] 在一些实施方案中，本公开的dsRNA分子包含交替模式的修饰，尤其在B1、B2、B3、B1'、B2'、B3'、B4'区域中。如本文中所用，术语"交替基序"或"交替模式"是指具有一个或多个修饰的基序，每一修饰出现在一条链的交替核苷酸上。交替核苷酸可是指每两个核苷酸一个或每三个核苷酸一个或类似模式。例如，如果A、B及C各自表示一种类型的对核苷酸的修饰，则交替基序可以是"ABABABABABAB…"、"AABBAABBAABB…"、"AABAABAABAAB…"、"AAABAAABAAAB…"、"AAABBBAAABBB…"或"ABCABCABCABC…"等。

[0519] 交替基序中含有的修饰的类型可以相同或不同。例如，如果A、B、C、D各自表示一种类型的对核苷酸的修饰，则交替基序即每两个核苷酸上的修饰可以是相同的，但正义链或反义链可各自选自交替基序如"ABABAB…"、"ACACAC…"、"BDBDBD…"或"CDCDCD…"等中修

饰的若干可能性。

[0520] 在一些实施方案中，本公开的dsRNA分子包含，正义链的交替基序的修饰模式相对于该反义链的交替基序的修饰模式变动。该变动可使得正义链的核苷酸的修饰基团对应于反义链的核苷酸的不同修饰基团，反之亦然。例如，当正义链与反义链在dsRNA双链体中碱基配对时，在该双链体区域内，正义链中的交替基序可始于该链的5'-3'的"ABABAB"，且反义链中的交替基序可始于该链的3'-5'的"BABABA"。作为另一示例，在双链体区域内，正义链中的交替基序可始于该链的5'-3'的"AABBAABB"，且反义链中的交替基序可始于该链的3'-5'的"BBAABBAA"，因此正义链与反义链之间存在修饰模式的完全或部分变动。

[0521] 本公开的dsRNA分子可进一步包含至少一个硫代磷酸酯或甲基膦酸酯核苷酸间键。硫代磷酸酯或甲基膦酸酯核苷酸间键修饰可出现在正义链或反义链或两者的位于该链任何位置的任何核苷酸上。例如，核苷酸间键修饰可出现在正义链和/或反义链的每一个核苷酸；每一核苷酸间键修饰可以交替模式出现在正义链或反义链；或正义链或反义链包含交替模式的两种核苷酸间键修饰。正义链的交替模式的核苷酸间键修饰可与反义链相同或不同，其正义链的交替模式的核苷酸间键修饰可具有相对于反义链的交替模式的核苷酸间键修饰的变动。

[0522] 在一些实施方案中，dsRNA分子包含位于突出端区域内的硫代磷酸酯或甲基膦酸酯核苷酸间键修饰。例如，突出端区域包含两个核苷酸且在这两个核苷酸间具有硫代磷酸酯或甲基膦酸酯核苷酸间键。核苷酸间键修饰还可制成以将该突出端核苷酸与双链体区域内的末端配对核苷酸键。例如，至少2、3、4或全部突出端核苷酸可通过硫代磷酸酯或甲基膦酸酯核苷酸间键而键联，且任选的，可存在将突出端核苷酸与作为该突出端核苷酸的下一个成对核苷酸连接的的额外的硫代磷酸酯或甲基膦酸酯核苷酸间键。例如，在末端三个核苷酸之间可能存在至少两个硫代硫酸酯核苷酸间键，其中该三个核苷酸中的两者是突出端核苷酸，且第三个核苷酸是紧邻该突出端核苷酸的下一个配对核苷酸。优选地，末端三个核苷酸可位于反义链的3'端。

[0523] 在一些实施方案中，dsRNA分子的正义链包含通过1、2、3、4、5、6、7、8、9、10、11、12、13、14、15或16个磷酸酯核苷酸间键分隔的1至10个具有两个或十个硫代磷酸酯或甲基膦酸酯核苷酸键的嵌段，其中该硫代磷酸酯或甲基膦酸酯核苷酸间键中的一个被置于该寡核苷酸序列中的任何位置，并且该正义链与包含硫代硫酸酯、甲基膦酸酯及磷酸酯核苷酸间键的任何组合的反义链或包含硫代磷酸酯或甲基膦酸酯或磷酸酯键的反义链配对。

[0524] 在一些实施方案中，dsRNA分子的反义链包含通过1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17或18个磷酸酯核苷酸间键分隔的2个具有两个硫代磷酸酯或甲基膦酸酯核苷酸键的嵌段，其中该硫代磷酸酯或甲基膦酸酯核苷酸间键中的一个被置于该寡核苷酸序列中的任何位置，并且该反义链与包含硫代硫酸酯、甲基膦酸酯及磷酸酯核苷酸间键的任何组合的正义链或包含硫代磷酸酯或甲基膦酸酯或磷酸酯键的正义链配对。

[0525] 在一些实施方案中，dsRNA分子的反义链包含通过1、2、3、4、5、6、7、8、9、10、11、12、13、14、15或16个磷酸酯核苷酸间键分隔的2个具有三个硫代磷酸酯或甲基膦酸酯核苷酸键的嵌段，其中该硫代磷酸酯或甲基膦酸酯核苷酸间键中的一个被置于该寡核苷酸序列中的任何位置，并且该反义链与包含硫代硫酸酯、甲基膦酸酯及磷酸酯核苷酸间键的任何组合的正义链或包含硫代磷酸酯或甲基膦酸酯或磷酸酯键的正义链配对。

[0526] 在一些实施方案中，dsRNA分子的反义链包含通过1、2、3、4、5、6、7、8、9、10、11、12、13或14个磷酸酯核苷酸间键分隔的2个具有四个硫代磷酸酯或甲基膦酸酯核苷酸键的嵌段，其中该硫代磷酸酯或甲基膦酸酯核苷酸间键中的一个被置于该寡核苷酸序列中的任何位置，并且该反义链与包含硫代硫酸酯、甲基膦酸酯及磷酸酯核苷酸间键的任何组合的正义链或包含硫代磷酸酯或甲基膦酸酯或磷酸酯键的正义链配对。

[0527] 在一些实施方案中，dsRNA分子的反义链包含通过1、2、3、4、5、6、7、8、9、10、11或12个磷酸酯核苷酸间键分隔的2个具有五个硫代磷酸酯或甲基膦酸酯核苷酸键的嵌段，其中该硫代磷酸酯或甲基膦酸酯核苷酸间键中的一个被置于该寡核苷酸序列中的任何位置，并且所述反义链与包含硫代硫酸酯、甲基膦酸酯及磷酸酯核苷酸间键的任何组合的正义链或包含硫代磷酸酯或甲基膦酸酯或磷酸酯键的正义链配对。

[0528] 在一些实施方案中，dsRNA分子的反义链包含通过1、2、3、4、5、6、7、8、9或10个磷酸酯核苷酸间键分隔的2个具有六个硫代磷酸酯或甲基膦酸酯核苷酸键的嵌段，其中该硫代磷酸酯或甲基膦酸酯核苷酸间键中的一个被置于该寡核苷酸序列中的任何位置，并且该反义链与包含硫代硫酸酯、甲基膦酸酯及磷酸酯核苷酸间键的任何组合的正义链或包含硫代磷酸酯或甲基膦酸酯或磷酸酯键的正义链配对。

[0529] 在一些实施方案中，dsRNA分子的反义链包含通过1、2、3、4、5、6、7或8个磷酸酯核苷酸间键分隔的2个具有七个硫代磷酸酯或甲基膦酸酯核苷酸键的嵌段，其中该硫代磷酸酯或甲基膦酸酯核苷酸间键中的一个被置于该寡核苷酸序列中的任何位置，并且该反义链与包含硫代硫酸酯、甲基膦酸酯及磷酸酯核苷酸间键的任何组合的正义链或包含硫代磷酸酯或甲基膦酸酯或磷酸酯键的正义链配对。

[0530] 在一些实施方案中，dsRNA分子的反义链包含通过1、2、3、4、5或6个磷酸酯核苷酸间键分隔的2个具有八个硫代磷酸酯或甲基膦酸酯核苷酸键的嵌段，其中该硫代磷酸酯或甲基膦酸酯核苷酸间键中的一个被置于该寡核苷酸序列中的任何位置，并且该反义链与包含硫代硫酸酯、甲基膦酸酯及磷酸酯核苷酸间键的任何组合的正义链或包含硫代磷酸酯或甲基膦酸酯或磷酸酯键的正义链配对。

[0531] 在一些实施方案中，dsRNA分子的反义链包含通过1、2、3或4个磷酸酯核苷酸间键分隔的2个具有九个硫代磷酸酯或甲基膦酸酯核苷酸键的嵌段，其中该硫代磷酸酯或甲基膦酸酯核苷酸间键中的一个被置于该寡核苷酸序列中的任何位置，并且该反义链与包含硫代硫酸酯、甲基膦酸酯及磷酸酯核苷酸间键的任何组合的正义链或包含硫代磷酸酯或甲基膦酸酯或磷酸酯键的正义链配对。

[0532] 在一些实施方案中，本公开的dsRNA分子进一步包含位于正义链或反义链的端位置1至10内的一个或多个硫代磷酸酯或甲基膦酸酯核苷酸间键。例如，至少2、3、4、5、6、7、8、9或10个核苷酸可通过硫代磷酸酯或甲基膦酸酯核苷酸间键而与正义链或反义链的一端或两端连接。

[0533] 在一些实施方案中，本公开的dsRNA分子进一步包含位于正义链或反义链各自的双链体内部区域1至10内的一个或多个硫代磷酸酯或甲基膦酸酯核苷酸间键。例如，至少2、3、4、5、6、7、8、9或10个核苷酸可通过硫代磷酸酯或甲基膦酸酯核苷酸间键而连接于自正义链的5’端计数的双链体区域的位置8至16处；dsRNA分子可任选的进一步包含位于端位置1至10内的一个或多个硫代磷酸酯或甲基膦酸酯核苷酸间键修饰。

[0534] 在一些实施方案中，本公开的dsRNA分子进一步包含位于正义链(自5’-端计数)的位置1至5内的一至五个硫代磷酸酯或甲基膦酸酯核苷酸间键修饰及位于位置18至23内的一至五个硫代磷酸酯或甲基膦酸酯核苷酸间键修饰，以及位于反义链(自5’端计数)的位置1及2处的一至五个硫代磷酸酯或甲基膦酸酯核苷酸间键修饰及位于位置18至23内的一至五个硫代磷酸酯或甲基膦酸酯核苷酸间键修饰。

[0535] 在一些实施方案中，本公开的dsRNA分子进一步包含位于正义链(自5’-端计数)的位置1至5内的一个硫代磷酸酯核苷酸间键修饰及位于位置18至23内的一个硫代磷酸酯或甲基膦酸酯核苷酸间键修饰，以及位于反义链(自5’端计数)的位置1及2处的一个硫代磷酸酯核苷酸间键修饰及位于位置18至23内的两个硫代磷酸酯或甲基膦酸酯核苷酸间键修饰。

[0536] 在一些实施方案中，本公开的dsRNA分子进一步包含位于正义链(自5’-端计数)的位置1至5内的两个硫代磷酸酯核苷酸间键修饰及位于位置18至23内的一个硫代磷酸酯核苷酸间键修饰，以及位于反义链(自5’端计数)的位置1及2处的一个硫代磷酸酯核苷酸间键修饰及位于位置18至23内的两个硫代磷酸酯核苷酸间键修饰。

[0537] 在一些实施方案中，本公开的dsRNA分子进一步包含位于正义链(自5’-端计数)的位置1至5内的两个硫代磷酸酯核苷酸间键修饰及位于位置18至23内的两个硫代磷酸酯核苷酸间键修饰，以及位于反义链(自5’端计数)的位置1及2处的一个硫代磷酸酯核苷酸间键修饰及位于位置18至23内的两个硫代磷酸酯核苷酸间键修饰。

[0538] 在一些实施方案中，本公开的dsRNA分子进一步包含位于正义链(自5’-端计数)的位置1至5内的两个硫代磷酸酯核苷酸间键修饰及位于位置18至23内的两个硫代磷酸酯核苷酸间键修饰，以及位于反义链(自5’端计数)的位置1及2处的一个硫代磷酸酯核苷酸间键修饰及位于位置18至23内的一个硫代磷酸酯核苷酸间键修饰。

[0539] 在一些实施方案中，本公开的dsRNA分子进一步包含位于正义链(自5’-端计数)的位置1至5内的一个硫代磷酸酯核苷酸间键修饰及位于位置18至23内的一个硫代磷酸酯核苷酸间键修饰，以及位于反义链(自5’端计数)的位置1及2处的两个硫代磷酸酯核苷酸间键修饰及位于位置18至23内的两个硫代磷酸酯核苷酸间键修饰。

[0540] 在一些实施方案中，本公开的dsRNA分子进一步包含位于正义链(自5’-端计数)的位置1至5内的一个硫代磷酸酯核苷酸间键修饰及位于位置18至23内的一个硫代磷酸酯核苷酸间键修饰，以及位于反义链(自5’端计数)的位置1及2处的两个硫代磷酸酯核苷酸间键修饰及位于位置18至23内的一个硫代磷酸酯核苷酸间键修饰。

[0541] 在一些实施方案中，本公开的dsRNA分子进一步包含位于正义链(自5’-端计数)的位置1至5内的一个硫代磷酸酯核苷酸间键修饰，以及位于反义链(自5’端计数)的位置1及2处的两个硫代磷酸酯核苷酸间键修饰及位于位置18至23内的一个硫代磷酸酯核苷酸间键修饰。

[0542] 在一些实施方案中，本公开的dsRNA分子进一步包含位于正义链(自5’-端计数)的位置1至5内的两个硫代磷酸酯核苷酸间键修饰，以及位于反义链(自5’端计数)的位置1及2处的一个硫代磷酸酯核苷酸间键修饰及位于位置18至23内的两个硫代磷酸酯核苷酸间键修饰。

[0543] 在一些实施方案中，本公开的dsRNA分子进一步包含位于正义链(自5’-端计数)的位置1至5内的两个硫代磷酸酯核苷酸间键修饰及位于位置18至23内的一个硫代磷酸酯核

苷酸间键修饰,以及位于反义链(自5’端计数)的位置1及2处的两个硫代磷酸酯核苷酸间键修饰及位于位置18至23内的一个硫代磷酸酯核苷酸间键修饰。

[0544] 在一些实施方案中,本公开的dsRNA分子进一步包含位于正义链(自5’-端计数)的位置1至5内的两个硫代磷酸酯核苷酸间键修饰及位于位置18至23内的一个硫代磷酸酯核苷酸间键修饰,以及位于反义链(自5’端计数)的位置1及2处的两个硫代磷酸酯核苷酸间键修饰及位于位置18至23内的两个硫代磷酸酯核苷酸间键修饰。

[0545] 在一些实施方案中,本公开的dsRNA分子进一步包含位于正义链(自5’-端计数)的位置1至5内的两个硫代磷酸酯核苷酸间键修饰及位于位置18至23内的一个硫代磷酸酯核苷酸间键修饰,以及位于反义链(自5’端计数)的位置1及2处的一个硫代磷酸酯核苷酸间键修饰及位于位置18至23内的两个硫代磷酸酯核苷酸间键修饰。

[0546] 在一些实施方案中,本公开的dsRNA分子进一步包含位于正义链(自5’-端计数)的位置1及2处的两个硫代磷酸酯核苷酸间键修饰及位于位置20及21处的两个硫代磷酸酯核苷酸间键修饰,以及位于反义链(自5’端计数)的位置1处的一个硫代磷酸酯核苷酸间键修饰及位于位置21处的一个硫代磷酸酯核苷酸间键修饰。

[0547] 在一些实施方案中,本公开的dsRNA分子进一步包含位于正义链(自5’-端计数)的位置1处的一个硫代磷酸酯核苷酸间键修饰及位于位置21处的一个硫代磷酸酯核苷酸间键修饰,以及位于反义链(自5’端计数)的位置1及2处的两个硫代磷酸酯核苷酸间键修饰及位于位置20及21处的两个硫代磷酸酯核苷酸间键修饰。

[0548] 在一些实施方案中,本公开的dsRNA分子进一步包含位于正义链(自5’-端计数)的位置1及2处的两个硫代磷酸酯核苷酸间键修饰及位于位置21及22处的两个硫代磷酸酯核苷酸间键修饰,以及位于反义链(自5’端计数)的位置1处的一个硫代磷酸酯核苷酸间键修饰及位于位置21处的一个硫代磷酸酯核苷酸间键修饰。

[0549] 在一些实施方案中,本公开的dsRNA分子进一步包含位于正义链(自5’-端计数)的位置1处的一个硫代磷酸酯核苷酸间键修饰及位于位置21处的一个硫代磷酸酯核苷酸间键修饰,以及位于反义链(自5’端计数)的位置1及2处的两个硫代磷酸酯核苷酸间键修饰及位于位置21及22处的两个硫代磷酸酯核苷酸间键修饰。

[0550] 在一些实施方案中,本公开的dsRNA分子进一步包含位于正义链(自5’-端计数)的位置1及2处的两个硫代磷酸酯核苷酸间键修饰及位于位置22及23处的两个硫代磷酸酯核苷酸间键修饰,以及位于反义链(自5’端计数)的位置1处的一个硫代磷酸酯核苷酸间键修饰及位于位置21处的一个硫代磷酸酯核苷酸间键修饰。

[0551] 在一些实施方案中,本公开的dsRNA分子进一步包含位于正义链(自5’-端计数)的位置1处的一个硫代磷酸酯核苷酸间键修饰及位于位置21处的一个硫代磷酸酯核苷酸间键修饰,以及位于反义链(自5’端计数)的位置1及2处的两个硫代磷酸酯核苷酸间键修饰及位于位置23及23处的两个硫代磷酸酯核苷酸间键修饰。

[0552] 在一些实施方案中,本公开的化合物包含骨架手性中心(backbone chiral center)的模式。在一些实施方案中,骨架手性中心的共有模式(common pattern)包含至少5个Sp构型的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含至少6个Sp构型的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含至少7个Sp构型的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含至少8个Sp构型的核苷酸间键。

在一些实施方案中,骨架手性中心的共有模式包含至少9个Sp构型的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含至少10个Sp构型的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含至少11个Sp构型的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含至少12个Sp构型的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含至少13个Sp构型的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含至少14个Sp构型的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含至少15个Sp构型的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含至少16个Sp构型的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含至少17个Sp构型的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含至少18个Sp构型的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含至少19个Sp构型的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含不超过8个Rp构型的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含不超过7个Rp构型的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含不超过6个Rp构型的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含不超过5个Rp构型的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含不超过4个Rp构型的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含不超过3个Rp构型的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含不超过2个Rp构型的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含不超过1个Rp构型的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含不超过8个非手性的核苷酸间键(作为非限制性示例,磷酸二酯)。在一些实施方案中,骨架手性中心的共有模式包含不超过7个非手性的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含不超过6个非手性的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含不超过5个非手性的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含不超过4个非手性的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含不超过3个非手性的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含不超过2个非手性的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含不超过1个非手性的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含至少10个Sp构型的核苷酸间键,以及不超过8个非手性的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含至少11个Sp构型的核苷酸间键,以及不超过7个非手性的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含至少12个Sp构型的核苷酸间键,以及不超过6个非手性的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含至少13个Sp构型的核苷酸间键,以及不超过6个非手性的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含至少14个Sp构型的核苷酸间键,以及不超过5个非手性的核苷酸间键。在一些实施方案中,骨架手性中心的共有模式包含至少15个Sp构型的核苷酸间键,以及不超过4个非手性的核苷酸间键。在一些实施方案中,Sp构型的核苷酸间键任选的是连续的或非连续的。在一些实施方案中,Rp构型的核苷酸间键任选的是连续的或非连续的。在一些实施方案中,非手性的核苷酸间键任选的是连续的或非连续的。

[0553] 在一些实施方案中,本公开的化合物包含嵌段,该嵌段是立体化学嵌段。在一些实施方案中,嵌段是Rp嵌段,其中该嵌段的每个核苷酸间键是Rp。在一些实施方案中,5'-嵌段是Rp嵌段。在一些实施方案中,3'-嵌段是Rp嵌段。在一些实施方案中,嵌段是Sp嵌段,其中

该嵌段的每个核苷酸间键是Sp。在一些实施方案中,5'-嵌段是Sp嵌段。在一些实施方案中,3'-嵌段是Sp嵌段。在一些实施方案中,所提供的寡核苷酸包含Rp嵌段及Sp嵌段两者。在一些实施方案中,所提供的寡核苷酸包含一个或多个Rp嵌段但不包含Sp嵌段。在一些实施方案中,所提供的寡核苷酸包含一个或多个Sp嵌段但不包含Rp嵌段。在一些实施方案中,所提供的寡核苷酸包含一个或多个PO嵌段,其中每个核苷酸间键是天然磷酸酯键。

[0554] 在一些实施方案中,本公开的化合物包含5'-嵌段,该嵌段是Sp嵌段,其中每个糖部分包含2'-F修饰。在一些实施方案中,5'-嵌段是Sp嵌段,其中每个核苷酸间键是经修饰的核苷酸间键,并且每个糖部分包含2'-F修饰。在一些实施方案中,5'-嵌段是Sp嵌段,其中每个核苷酸间键是硫代磷酸酯键,并且每个糖部分包含2'-F修饰。在一些实施方案中,5'-嵌段包含4个或更多个核苷酸单元。在一些实施方案中,5'-嵌段包含5个或更多个核苷酸单元。在一些实施方案中,5'-嵌段包含6个或更多个核苷酸单元。在一些实施方案中,5'-嵌段包含7个或更多个核苷酸单元。在一些实施方案中,3'-嵌段是Sp嵌段,其中每个糖部分包含2'-F修饰。在一些实施方案中,3'-嵌段是Sp嵌段,其中每个核苷酸间键是经修饰的核苷酸间键,并且每个糖部分包含2'-F修饰。在一些实施方案中,3'-嵌段是Sp嵌段,其中每个核苷酸间键是硫代磷酸酯键,并且每个糖部分包含2'-F修饰。在一些实施方案中,3'-嵌段包含4个或更多个核苷酸单元。在一些实施方案中,3'-嵌段包含5个或更多个核苷酸单元。在一些实施方案中,3'-嵌段包含6个或更多个核苷酸单元。在一些实施方案中,3'-嵌段包含7个或更多个核苷酸单元。

[0555] 在一些实施方案中,本公开的化合物包含位于一区域中的一种类型的核苷酸或寡核苷酸,其后为特定类型的核苷酸间键,例如,天然磷酸酯键、经修饰的核苷酸间键、Rp手性核苷酸间键、Sp手性核苷酸间键等。在一些实施方案中,A之后是Sp。在一些实施方案中,A之后是Rp。在一些实施方案中,A之后是天然磷酸酯键(PO)。在一些实施方案中,U之后是Sp。在一些实施方案中,U之后是Rp。在一些实施方案中,U之后是天然磷酸酯键(PO)。在一些实施方案中,C之后是Sp。在一些实施方案中,C之后是Rp。在一些实施方案中,C之后是天然磷酸酯键(PO)。在一些实施方案中,G之后是Sp。在一些实施方案中,G之后是Rp。在一些实施方案中,G之后是天然磷酸酯键(PO)。在一些实施方案中,C及U之后是Sp。在一些实施方案中,C及U之后是Rp。在一些实施方案中,C及U之后是天然磷酸酯键(PO)。在一些实施方案中,A及G之后是Sp。在一些实施方案中,A及G之后是Rp。

[0556] 在一些实施方案中,反义链包含位于核苷酸位置21与22之间以及核苷酸位置22与23之间的硫代磷酸酯核苷酸间键,其中反义链含有定位在反义链的种子区域内(即,位于反义链的5'端的位置2至9处)的双链体的至少一个热去稳定化修饰,并且其中dsRNA任选的进一步具有下述特征的至少一个(例如,一个、两个、三个、四个、五个、六个、七个或全部八个):(i)反义链包含2、3、4、5或6个2'-氟修饰;(ii)反义链包含3、4或5个硫代磷酸酯核苷酸间键;(iii)正义链与配体缀合;(iv)正义链包含2、3、4或5个2'-氟修饰;(v)正义链包含1、2、3、4或5个硫代磷酸酯核苷酸间键;(vi)dsRNA包含至少四个2'-氟修饰;(vii)dsRNA包含长度为12至40个核苷酸对的双链体区域;以及(viii)dsRNA具有位于反义链5'端的钝端。

[0557] 在一些实施方案中,反义链包含位于核苷酸位置1与2之间、核苷酸位置2与3之间、核苷酸位置21与22之间以及核苷酸位置22与23之间的硫代磷酸酯核苷酸间键,其中反义链含有定位在反义链的种子区域内(即,位于反义链的5'端的位置2至9处)的双链体的至少一

个热去稳定化修饰，并且其中dsRNA任选的进一步具有下述特征的至少一个(例如，一个、两个、三个、四个、五个、六个、七个或全部八个)：(i)反义链包含2、3、4、5或6个2'-氟修饰；(ii)正义链与配体缀合；(iii)正义链包含2、3、4或5个2'-氟修饰；(iv)正义链包含1、2、3、4或5个硫代磷酸酯核苷酸间键；(v)dsRNA包含至少四个2'-氟修饰；(vi)dsRNA包含长度为12至40个核苷酸对的双链体区域；(vii)dsRNA包含长度为12至40个核苷酸对的双链体区域；以及(viii)dsRNA具有位于反义链5'端的钝端。

[0558] 在一些实施方案中，正义链包含位于核苷酸位置1与2之间以及核苷酸位置2与3之间的硫代磷酸酯核苷酸间键，其中反义链含有定位在反义链的种子区域内(即，位于反义链的5'端的位置2至9处)的双链体的至少一个热去稳定化修饰，并且其中dsRNA任选的进一步具有下述特征的至少一个(例如，一个、两个、三个、四个、五个、六个、七个或全部八个)：(i)反义链包含2、3、4、5或6个2'-氟修饰；(ii)反义链包含1、2、3、4或5个硫代磷酸酯核苷酸间键；(iii)正义链与配体缀合；(iv)正义链包含2、3、4或5个2'-氟修饰；(v)正义链包含3、4或5个硫代磷酸酯核苷酸间键；(vi)dsRNA包含至少四个2'-氟修饰；(vii)dsRNA包含长度为12至40个核苷酸对的双链体区域；以及(viii)dsRNA具有位于反义链5'端的钝端。

[0559] 在一些实施方案中，正义链包含位于核苷酸位置1与2之间以及核苷酸位置2与3之间的硫代磷酸酯核苷酸间键，反义链包含位于核苷酸位置1与2之间、核苷酸位置2与3之间、核苷酸位置21与22之间以及核苷酸位置22与23之间的硫代磷酸酯核苷酸间键，其中反义链含有定位在反义链的种子区域内(即，位于反义链的5'端的位置2至9处)的双链体的至少一个热去稳定化修饰，并且其中dsRNA任选的进一步具有下述特征的至少一个(例如，一个、两个、三个、四个、五个、六个或全部七个)：(i)反义链包含2、3、4、5或6个2'-氟修饰；(ii)正义链与配体缀合；(iii)正义链包含2、3、4或5个2'-氟修饰；(iv)正义链包含3、4或5个硫代磷酸酯核苷酸间键；(v)dsRNA包含至少四个2'-氟修饰；(vi)dsRNA包含长度为12至40个核苷酸对的双链体区域；以及；(vii)dsRNA具有位于反义链5'端的钝端。

[0560] 在一些实施方案中，本公开的dsRNA分子包含与靶标的错配、双链体中的错配、或其组合。错配可出现在突出端区域内或双链体区域内。基于碱基对促进解离或熔融的倾向性(例如，基于特定配对的关联或解离的自由能，自由能是基于单个配对检查配对的最简单的途径，但还可使用次近邻分析及类似分析)，可将碱基对排名。就促进解离而言：A:U优于G:C；G:U优于G:C；且I:C优于G:C(I是肌苷)。错配例如非标准配对或除标准配对之外的(如本文中他处所描述)优于标准(A:T、A:U、G:C)配对；且包括万用碱基的配对优于标准配对。

[0561] 在一些实施方案中，本公开的dsRNA分子包含，位于该双链体区域内的反义链中从5'端计数最前列的第1、2、3、4或5个碱基对的至少一个可独立选自下列所组成的组：A:U、G:U、I:C、以及错配(例如非标准配对或除标准配对之外的配对或包括万用碱基的配对)，以促进反义链于该双链体的5'端的解离。

[0562] 在一些实施方案中，位于该双链体区域内的反义链中从5'端计数的第1个位置处的核苷酸是选自由A、dA、dU、U及dT所组成的组。另选地，位于该双链体区域内的反义链中从5'末端计数最前列的第1、2或3个碱基对的至少一个是AU碱基对。例如，位于该双链体区域内的反义链中从5'端计数的第一个碱基对是AU碱基对。

[0563] 发现，将4'-修饰或5'-修饰的核苷酸引入位于单链或双链寡核苷酸任何位置处的二核苷酸的磷酸二酯(PO)、硫代磷酸酯(PS)或二硫代磷酸酯(PS2)键的3'-末端，可对该核

苷酸间键发挥立体效应,并依次保护或稳定化该键以对抗核酸酶。

[0564] 在一些实施方案中,将5'-修饰的核苷引入位于单链或双链siRNA任何位置处的二核苷酸的3'端。例如,可将5'-烷基化的核苷引入位于单链或双链siRNA任何位置处的二核苷酸的3'末端。位于核糖的5'位置处的烷基基团可以是外消旋物或手性纯R或S异构体。示例性5'-烷基化的核苷是5'-甲基核苷。5'-甲基可以是外消旋物或手性纯R或S异构体。

[0565] 在一些实施方案中,将4'-修饰的核苷引入位于单链或双链siRNA任何位置处的二核苷酸的3'末端。例如,可将4'-烷基化的核苷引入位于单链或双链siRNA任何位置处的二核苷酸的3'端。位于核糖的4'位置处的烷基基团可以是外消旋物或手性纯R或S异构体。示例性4'-烷基化的核苷是4'-甲基核苷。4'-甲基可以是外消旋物或手性纯R或S异构体。另选地,可将4'-O-烷基化的核苷引入位于单链或双链siRNA任何位置处的二核苷酸的3'端。核糖的4'-O-烷基可以是外消旋物或手性纯R或S异构体。示例性4'-O-烷基化的核苷是4'-O-甲基核苷。4'-O-甲基可以是外消旋物或手性纯R或S异构体。

[0566] 在一些实施方案中,将5'-烷基化的核苷引入dsRNA的正义链或反义链的任何位置处,并且此类修饰维持或改进dsRNA的效力。5'-烷基可以是外消旋物或手性纯R或S异构体。示例性5'-烷基化的核苷是5'-甲基核苷。5'-甲基可以是外消旋物或手性纯R或S异构体。

[0567] 在一些实施方案中,将4'-烷基化的核苷引入dsRNA的正义链或反义链的任何位置处,并且此类修饰维持或改进dsRNA的效力。4'-烷基可以是外消旋物或手性纯R或S异构体。示例性4'-烷基化的核苷是4'-甲基核苷。4'-甲基可以是外消旋物或手性纯R或S异构体。

[0568] 在一些实施方案中,将4'-O-烷基化的核苷引入dsRNA的正义链或反义链的任何位置处,并且此类修饰维持或改进dsRNA的效力。5'-烷基可以是外消旋物或手性纯R或S异构体。示例性4'-O-烷基化的核苷是4'-O-甲基核苷。4'-O-甲基可以是外消旋物或手性纯R或S异构体。

[0569] 在一些实施方案中,本公开的dsRNA分子可包含2'-5'键(与2'-H、2'-OH及2'-OMe键以及与P=O或P=S键)。例如,2'-5'键修饰可用来促进核酸酶抗性或用来抑制正义链与反义链的结合,或可用于正义链的5'末端以避免正义链被RISC激活。

[0570] 在另一实施方案中,本公开的dsRNA分子可包含L-糖(例如,L-核糖、L-阿拉伯糖,其具有2'-H、2'-OH及2'-OMe)。例如,这些L-糖可用来促进核酸酶抗性或用来抑制正义链与反义链的结合,或可用于正义链的5'末端以避免正义链被RISC激活。

[0571] 多个出版物描述可用于本公开的dsRNA中的多聚siRNA。此类出版物是包括WO2007/091269、US 7858769、WO2010/141511、WO2007/117686、WO2009/014887及WO2011/031520,其各自通过援引而以其整体并入本文。

[0572] 如下文中更详细描述,含有一个或多个碳水化合物部分与RNAi剂的缀合的RNAi剂可优化该RNAi剂的一种或多种特性。在多种情形中,碳水化合物部分将附接至该RNAi剂的经修饰的子单元。例如,dsRNA剂的一个或多个核糖核苷酸子单元的核糖可替换为另一部分,如其上附接有碳水化合物配体的非碳水化合物(优选是环状)载剂。本文中,其子单元的核糖已经由此替换的核糖核苷酸子单元称为核糖替换修饰子单元(RRMS)。环状载剂可以是碳环系统,即,所有环原子都是碳原子;或是杂环系统,即,一个或多个环原子可以是杂原子如氮、氧、硫。环状载剂可以是单环系统,或可含有两个或多个环如稠环。环状载剂可以是完全饱和的环系统,或其可含有一个或多个双键。

[0573] 配体可经由载剂附接至多核苷酸。载剂包括(i)至少一个“骨架附接点”,优选两个“骨架附接点”,以及(ii)至少一个“系带附接点”。如本文中所用,“骨架附接点”是指官能基团如羟基,或通常是键,其可用于且适用于将该载剂并入核糖核酸的骨架(如磷酸酯或经修饰的磷酸酯(如含硫的)中。在一些实施方案中,“系带附接点”(TAP)是指环状载剂的构成环原子,例如,碳原子或杂原子(与提供骨架附接点的原子截然不同),其连接所选择的部分。该部分可以是例如碳水化合物,如单糖、二糖、三糖、四糖、寡糖及多糖。任选的,所选择的部分通过中介系带连接至所选择的载剂。因此,环状载剂一般将包括官能基团例如氨基,或通常提供适用于将另一化学实体如配体并入或系带至构建环的键。

[0574] RNAi剂可经由载剂缀合至配体,其中载剂可以是环状基团或非环状基团;优选地,环状基团是选自吡咯烷基、吡唑啉基、吡唑烷基、咪唑啉基、咪唑烷基、哌啶基、哌嗪基、[1,3]二氧杂环戊基、噁唑烷基、异噁唑烷基、吗啉基、噻唑烷基、异噻唑烷基、喹喔啉基、哒嗪酮基、四氢呋喃基及十氢萘基;优选地,该非环状基团是选自丝氨醇骨架或二乙醇胺骨架。

[0575] 在某些具体实施方案中,用于本公开的方法中的RNAi剂是选自由表10A、10B、12至15、19或20中任一者中所列的剂所组成的组。这些剂可进一步包含配体。

[0576] IV.缀合至配体的iRNA

[0577] 本发明的iRNA的RNA的另一修饰涉及将一个或多个增强该iRNA的活性、细胞分布或细胞摄取的配体、部分或缀合物化学键联至该iRNA。这些部分包括但不限于脂质部分如胆固醇部分(Letsinger et al.,Proc.Natl.Acid.Sci.USA,1989,86:6553-6556)、胆酸(Manoharan et al.,Biorg.Med.Chem.Let.,1994,4:1053-1060)、硫醚例如己基-S-三苯甲基硫醇(Manoharan et al.,Ann.N.Y.Acad.Sci.,1992,660:306-309;Manoharan et al.,Biorg.Med.Chem.Let.,1993,3:2765-2770)、硫代胆固醇(Oberhauser et al.,Nucl.Acids Res.,1992,20:533-538)、脂肪链例如十二烷二醇或十一烷基残基(Saison-Behmoaras et al.,EMBO J,1991,10:1111-1118;Kabanov et al.,FEBS Lett.,1990,259:327-330;Svinarchuk et al.,Biochimie,1993,75:49-54)、磷脂质例如二-十六烷基-rac-甘油或1,2-二-O-十六烷基-rac-甘油-3-磷酸三乙铵(Manoharan et al.,Tetrahedron Lett.,1995,36:3651-3654;Shea et al.,Nucl.Acids Res.,1990,18:3777-3783)、聚胺或聚乙二醇链(Manoharan et al.,Nucleosides&Nucleotides,1995,14:969-973)、或金刚烷乙酸(Manoharan et al.,Tetrahedron Lett.,1995,36:3651-3654)、棕榈酰基部分(Mishra et al.,Biochim.Biophys.Acta,1995,1264:229-237)、或十八烷基胺或己基胺-羰氧基胆固醇部分(Crooke et al.,J.Pharmacol.Exp.Ther.,1996,277:923-937)。

[0578] 在某些实施方案中,配体改变其所并入的iRNA剂的分布、靶向或寿命。在一些实施方案中,配体提供相比缺失此配体者增强的对于所选靶标如分子、细胞或细胞类型、区室如细胞或器官的区室、组织、器官或身体区域的亲和性。典型的配体将不会参与双链体核酸中的双链体配对。

[0579] 配体可包括天然出现的物质,例如蛋白质(例如,人血清白蛋白(HSA)、低密度脂蛋白(LDL)、或球蛋白);碳水化合物(例如,聚葡萄糖、聚三葡萄糖、几丁质、几丁聚糖、菊糖、环糊精或透明质酸);或脂质。配体还可以是重组分子或合成分子,诸如合成聚合物,例如合成聚氨基酸。聚氨基酸的示例包括聚赖氨酸(PLL)、聚L-天冬氨酸、聚L-谷氨酸、苯乙烯-马来酸酐共聚物、聚(L-乳酸交酯-共-乙交酯)共聚物、二乙烯基醚-马来酸酐共聚物、N-(2-羟基

丙基)甲基丙烯酰胺共聚物(HMPA)、聚乙二醇(PEG)、聚乙烯醇(PVA)、聚胺酯、聚(2-乙基丙烯酸)、N-异丙基丙烯酰胺聚合物、或聚磷嗪(polyphosphazine)。聚胺的示例包括:聚乙烯亚胺、聚赖氨酸(PLL)、精胺、精三胺、聚胺、假肽-聚胺、肽模拟性聚胺、树枝状聚胺、精氨酸、脒、鱼精蛋白、阳离子脂质、阳离子卟啉、聚胺的四级盐、或α-螺旋肽。

[0580] 配体还可包括靶向基团,例如细胞或组织靶向剂,例如凝集素、糖蛋白、脂质或蛋白质,例如抗体,其结合至特定的细胞类型如肾细胞。靶向基团可以是促甲状腺素、促黑素、凝集素、糖蛋白、表面活性剂蛋白A、黏蛋白碳水化合物、多价乳糖、多价半乳糖、N-乙酰基-半乳糖胺、N-乙酰基葡糖胺、多价甘露糖、多价果糖、糖基化聚氨基酸、多价半乳糖、运铁蛋白、双膦酸酯、聚谷氨酸盐、聚天冬氨酸盐、脂质、胆固醇、类固醇、胆汁酸、叶酸、维生素B12、生物素、或RGD肽或RGD肽模拟物。在某些实施方案中,配体是多价半乳糖,例如,N-乙酰基-半乳糖胺。

[0581] 配体的其他示例包括染料;嵌入剂(例如,吖啶类);交联剂(例如,补骨脂素、丝裂霉素C);卟啉类(TPPC4、texaphyrin、Sapphyrin);多环芳烃(例如,酚嗪(phenazine)、二氢酚嗪(dihydrophenazine));人工核酸内切酶(例如,EDTA);亲脂性分子(例如,胆固醇、胆酸、金刚烷乙酸、1-芘丁酸、双氢睾酮、1,3-双-O(十六烷基)甘油、香叶基氧己基、十六烷基甘油、冰片、薄荷醇、1,3-丙二醇、十六烷基、棕榈酸、肉豆蔻酸、O3-(油酰基)石胆酸、O3-(油酰基)胆烯酸、二甲氧基三苯甲基、或吩噁嗪);以及肽复合物(例如,触角足突变肽、Tat肽)、烷基化剂、磷酸盐、胺基、巯基、PEG(例如,PEG-40K)、MPEG、[MPEG]2、聚胺基、烷基、经取代的烷基、放射标记的标记物、酶、半抗原(如,生物素)、转运/吸收促进剂(例如,阿司匹林、维生素E、叶酸)、合成核糖核酸酶(例如,咪唑、双咪唑、组胺、咪唑簇、吖啶-咪唑缀合物、四氮杂大环的Eu3+复合物)、二硝基苯基、HRP、或AP。

[0582] 配体可以是蛋白质如糖蛋白、或肽如具有对于共配体的特异亲和性的分子、或抗体如结合至特定细胞类型诸如癌细胞、内皮细胞或骨细胞的抗体。配体还可包括激素及激素受体。它们还可包括非肽类物质,例如脂质、凝集素、碳水化合物、维生素、辅助因子、多价乳糖、多价半乳糖、N-乙酰基半乳糖胺、N-乙酰基葡糖胺、多价甘露糖、或多价果糖。配体可以是,例如,脂质多糖、p38 MAP激酶的活化剂、或NF-κB的活化剂。

[0583] 配体可以是例如药物的物质,其可通过例如扰乱细胞的细胞骨架如通过扰乱细胞的微管、微丝或中间体丝而增加细胞对iRNA剂的摄取。该药物可以是,例如,泰素(taxon)、长春新碱、长春碱、细胞松弛素、诺考达唑、茉莉素(japlakinolide)、红海海绵蛋白A、鬼笔环肽(phalloidin)、海洋苔藓素(swinholide)A、吲达诺欣(indanocine)、或肌基质蛋白(myoservin)。

[0584] 在一些实施方案中,附接至本文中所述的iRNA的配体用作药物动力学调节剂(PK调节剂)。PK调节剂是包括亲脂质类、胆汁酸、类固醇、磷脂质类似物、肽、蛋白质结合剂、PEG、维生素等。例示性PK调节剂是包括但不限于,胆固醇、脂肪酸、胆酸、石胆酸、二烷基甘油酯、二酰基甘油酯、磷脂质、神经脂质、萘普生(naproxen)、布洛芬(ibuprofen)、微生物E、生物素等。包含大量硫代硫酸酯键的寡核苷酸也已知结合至血清蛋白,因此在骨架中包含多个硫代磷酸酯键的短寡核苷酸如约5个碱基、10个碱基、15个碱基或20个碱基的寡核苷酸还可作为配体(例如,作为PK调节配体)而遵循本发明。此外,在本文所描述的实施方案中,结合血清组分(例如,血清蛋白)的适配体也适用于作为PK调节性配体而使用。

[0585] 本发明的缀合有配体的iRNA可通过使用携带侧链反应性官能度的寡核苷酸(如衍生自将连接分子附接在寡核苷酸(描述于下)上)而合成。这一反应性寡核苷酸可直接与可商购的配体、合成为携带多种保护基团的任一者的配体、或具有附接于其上的连接部分的配体反应。

[0586] 在本发明的缀合物中使用的寡核苷酸可通过熟知的固相合成技术便利且常规性地合成。用于此合成的设备可由多个供货商售卖,包括,例如,Applied Biosystems® (Foster City,Calif.)。可额外地或作为另一种选择地采用本领域中已知的用于此合成的任何其他手段。使用类似技术来制备其他寡核苷酸如硫代磷酸酯及烷基化衍生物也是已知的。

[0587] 在本发明的缀合有配体的寡核苷酸及携带序列特异性连接的核苷的配体分子中,该寡核苷酸及寡核苷可在合适DNA合成器上使用标准核苷酸或核苷前体、或已经携带连接部分的核苷酸或核苷缀合前体、已经携带配体分子的配体-核苷酸或核苷缀合前体、或携带配体的非核苷酸构建模块而组装。

[0588] 当使用已经携带连接部分的核苷酸缀合前体时,典型是完成序列特异性连接的核苷的合成,随后将该配体分子与该连接部分反应以形成缀合有配体的寡核苷酸。在一些实施方案中,本发明的寡核苷酸或连接的核苷是通过自动合成器使用除标准亚磷酰胺及可商购且常规用于寡核苷酸合成中的非标准亚磷酰胺之外的、衍生自配体-核苷缀合物的亚磷酰胺(phosphoramidite)来合成。

[0589] A.脂质缀合物

[0590] 在某些实施方案中,配体或缀合物是脂质或基于脂质的分子。此类脂质或基于脂质的分子典型可结合血清蛋白诸如人血清白蛋白(HSA)。HSA结合配体允许缀合物分布于靶标组织,如非肾脏的身体靶标组织。例如,靶标组织可以是肝脏,包括肝脏的实质细胞。可结合HSA的其他分子还可用作配体。例如,可使用萘普生或阿司匹林。脂质或基于脂质的配体可(a)增加对于缀合物降解的抗性,(b)增加对靶标细胞或细胞膜的靶向或递送,或(c)可用以调节与血清蛋白例如HSA的结合。

[0591] 基于脂质的配体可用以调节例如控制(例如,抑制)缀合物结合至靶标组织。例如,与HSA的结合更强的脂质或基于脂质的配体将更不可能靶向肾脏,并因此更不可能被从身体清除。与HSA的结合强度较低的脂质或基于脂质的配体可用来使缀合物靶向肾脏。

[0592] 在某些实施方案中,基于脂质的配体结合HSA。例如,配体可以足够的亲和性结合HSA,使得缀合物至非肾脏组织的分布得以增强。然而,亲和性的强度典型地不足以导致该HSA-配体的结合不可逆。

[0593] 在某些实施方案中,基于脂质的配体与HSA的结合弱或根本不结合,使得缀合物至肾脏的分布得以增强。靶向肾细胞的其他部分还可用于替换该基于脂质的配体或与该基于脂质的配体同时使用。

[0594] 在另一方面,配体是被靶标细胞例如增殖细胞摄取的部分例如维生素。这些是尤其可用于治疗以例如恶性或非恶性细胞如癌细胞的不期望的细胞增殖为特征的疾患。示例性维生素包括维生素A、维生素E及维生素K。其他示例性维生素包括B族维生素,例如叶酸、维生素B12、核黄素、生物素、吡哆醛、或其他被癌细胞摄取的维生素或营养物质。还包括HSA及低密度脂蛋白(LDL)。

[0595] B.细胞渗透剂

[0596] 在另一方面，该配体是细胞渗透剂，诸如螺旋细胞渗透剂。在某些实施方案中，该剂是两亲性剂。示例性剂是肽，例如tat或触角足突变肽。如果该剂是肽，其可经修饰，包括肽基模拟物、嵌入体、非肽或假肽类键、以及D-氨基酸的使用。螺旋剂典型是α-螺旋剂，并且可具有亲脂相及疏脂相。

[0597] 该配体可以是肽或肽模拟物。肽模拟物(本文中也称为寡肽模拟物)是能折叠为所定义的类似于天然肽的三维结构的分子。肽及肽模拟物至iRNA剂的附接可影响iRNA的药物动力学分布，例如通过提升细胞识别及吸收而影响。肽或肽模拟物部分可以是5至50个氨基酸的长度，例如，约5、10、15、20、25、30、35、40、45或50个氨基酸的长度。

[0598] 肽或肽模拟物可以是，例如，细胞渗透肽、阳离子肽、两亲性肽、或疏水性肽(例如，主要由Tyr、Trp或Phe构成)。肽部分可以是树枝状肽、受约束的肽或交联的肽。在另一选择中，肽部分可包括疏水性膜易位序列(MTS)。示例性的含有MTS的疏水性肽是具有下述氨基酸序列的RFGF:AAVALLPAVLLALLAP(SEQ ID NO:____)。含有疏水性MTS的RFGF类似物(例如，氨基酸序列AALLPVLLAAP(SEQ ID NO:____))还可以是靶向部分。肽部分可以是“递送性”肽，其可携带包括肽、寡核苷酸、及蛋白质在内的极性大分子跨越细胞膜。例如，已经发现，来自HIV Tat蛋白质的序列(GRKKRRQRRRPPQ(SEQ ID NO:____))及来自果蝇触角足突变肽蛋白的序列(RQIKIWFQNRRMKWKK(SEQ ID NO:____))能起到递送性肽的功能。肽或肽模拟物可由DNA的随机序列编码，例如从噬菌体展示文库或一珠一物(OBOC)组合库鉴定的肽(Lam et al.,Nature,354:82-84,1991)。典型地，经由合并的单体单元系带至dsRNA剂的肽或肽模拟物是细胞靶向肽诸如精氨酸-甘氨酸-天冬氨酸(RGD)肽或RGD模拟物。肽部分的长度范围可以是约5个氨基酸至约40个氨基酸。这些肽部分可具有结构性修饰，例如以增加稳定性或引导构象特性。可使用下文所述的任何结构性修饰。

[0599] 用于本发明的组合物及方法的RGD肽可以是线性或环状，且可经修饰例如经糖基化或甲基化以促进对特定组织的靶向。含有RGD的肽及肽模拟物可包括D-氨基酸，以及合成RGD模拟物。除了RGD的外，还可使用其他以整合素配体为靶标的部分。这一配体的优选缀合物靶向PECAM-1或VEGF。

[0600] RGD肽部分可用来靶向具体细胞类型，例如肿瘤细胞，诸如内皮瘤细胞或乳癌肿瘤细胞(Zitzmann et al.,Cancer Res.,62:5139-43,2002)。RGD肽可促成dsRNA剂靶向各种其他组织(包括肺、肾、脾或肝)的肿瘤(Aoki et al.,Cancer Gene Therapy8:783-787,2001)。典型地，RGD肽将促成iRNA剂靶向肾脏。RGD肽可以是线性或环状，并且可经修饰例如糖基化或甲基化以促成靶向特定组织。例如，经糖基化的RGD肽可将iRNA剂递送至表达αVβ3的肿瘤细胞(Haubner et al.,Jour.Nucl.Neuromolecular Med.42:326-336,2001)。

[0601] “细胞渗透肽”能渗透细胞例如微生物细胞诸如细菌或真菌细胞，或哺乳动物细胞诸如人类细胞。微生物细胞渗透肽可以是，例如，α-螺旋线性肽(例如，LL-37或Ceropin P1)、含二硫键的肽(例如，α-防御素、β-防御素或bactenecin)、或仅含有一个或两个支配性氨基酸的肽(例如，PR-39或indolicidin)。细胞渗透肽还可包括核定位信号(NLS)。例如，细胞渗透肽可以是双向两亲性肽如MPG，其是衍生自HIV-1gp41的融合肽结构域及SV40的T抗原的NLS(Simeoni et al.,Nucl.Acids Res.31:2717-2724,2003)。

[0602] C.碳水化合物缀合物

[0603] 在本发明的组合物及方法的一些实施方案中，iRNA进一步包含碳水化合物。缀合有碳水化合物的iRNA对于核酸的体内递送具有优势，且组合物是适用于体内治疗性用途，如本文中所描述。如本文中所用，“碳水化合物”是指碳水化合物自身，其由一个或多个具有至少6个碳原子的单糖单元(其可以是线性、分支或环状)与键合至每一碳原子的氧、氮或硫原子所组成；或是具有碳水化合物部分作为其一部分的化合物，该碳水化合物部分由一个或多个具有至少六个碳原子的单糖单元(其可以是线性、分支或环状)与键合至每一碳原子的氧、氮或硫原子所组成。代表性碳水化合物包括糖类(单糖、二糖、三糖及含有约4、5、6、7、8、或9个单糖单元的寡糖)，以及多糖如淀粉、糖原、纤维素及多糖胶。具体的单糖包括C5糖类及以上的(例如，C5、C6、C7或C8)糖类；二糖及三糖，其包括具有两个或三个单糖单元(例如，C5、C6、C7或C8)的糖类。

[0604] 在某些实施方案中，碳水化合物缀合物包含单糖。

[0605] 在某些实施方案中，单糖是N-乙酰基半乳糖胺(GalNAc)。GalNAc缀合物，其包含一种或多种N-乙酰基半乳糖胺(GalNAc)衍生物，描述于例如US 8,106,022中，该专利的全部内容通过援引而并入本文。在一些实施方案中，GalNAc缀合物用作配体，其使iRNA靶向具体细胞。在一些实施方案中，GalNAc缀合物使iRNA靶向肝脏细胞，例如，通过用作肝脏细胞(例如，肝细胞)的去唾液酸糖蛋白受体配体。

[0606] 在一些实施方案中，碳水化合物缀合物包含一个或多个GalNAc衍生物。GalNAc衍生物可经由接头(例如二价或三价分支接头)附接。在一些实施方案中，GalNAc缀合物缀合至正义链的3’端。在一些实施方案中，GalNAc缀合物经由接头例如本文所描述的接头缀合至iRNA剂(例如，缀合至正义链的3’端)。在一些实施方案中，GalNAc缀合物缀合至正义链的5’端。在一些实施方案中，GalNAc缀合物经由接头(例如本文所描述的接头)缀合至iRNA剂(例如，缀合至正义链的5’端)。

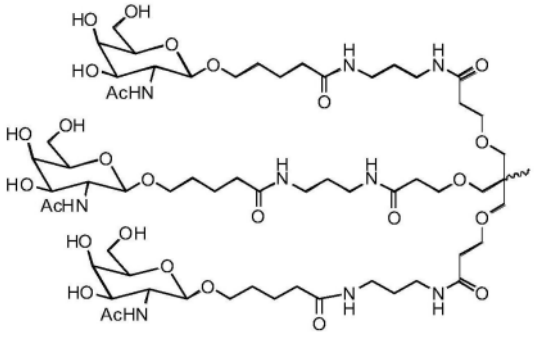
[0607] 在本发明的某些实施方案中，GalNAc或GalNAc衍生物经由单价接头附接至本发明的iRNA剂。在一些实施方案中，GalNAc或GalNAc衍生物经由二价接头附接至本发明的iRNA剂。在本发明的又一些实施方案中，GalNAc或GalNAc衍生物经由三价接头附接至本发明的iRNA剂。在本发明的其他实施方案中，GalNAc或GalNAc衍生物经由四价接头附接至本发明的iRNA剂。

[0608] 在某些实施方案中，本发明的双链RNAi剂包含一个附接至该iRNA剂的GalNAc或GalNAc衍生物。在某些实施方案中，本发明的双链RNAi剂含多(例如，2、3、4、5或6)个GalNAc或GalNAc衍生物，其通过多个单价接头而各自独立附接至该双链RNAi剂的多个核苷酸。

[0609] 在一些实施方案中，例如，当本发明的iRNA剂的两条链都是一个更大分子的一部分且通过位于一条链的3’端与另一条链的5’端之间的不间断核苷酸链连接而形成包含多个未配对核苷酸的发夹环时，该发夹环内的每一个未配对核苷酸可独立包含经由单价接头附接的GalNAc或GalNAc衍生物。发夹环还可通过双链体的一条链中的延伸突出形成。

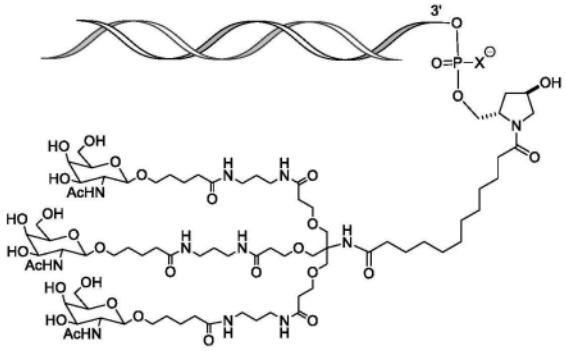
[0610] 在一些实施方案中，例如，当本发明的iRNA剂的两条链都是一个更大分子的一部分且通过位于一条链的3’端与另一条链的5’端之间的不间断核苷酸链连接而形成包含多个未配对核苷酸的发夹环时，该发夹环内的每一个未配对核苷酸可独立包含经由单价接头附接的GalNAc或GalNAc衍生物。发夹环还可通过双链体的一条链中的延伸突出端形成。

[0611] 在一些实施方案中，GalNAc缀合物是

[0612] 

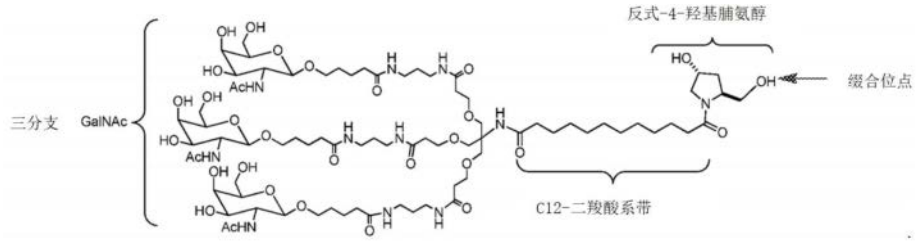
式 II。

[0613] 在一些实施方案中,RNAi剂经由下式所示的接头附接至碳水化合物缀合物,其中,X是O或S

[0614] 

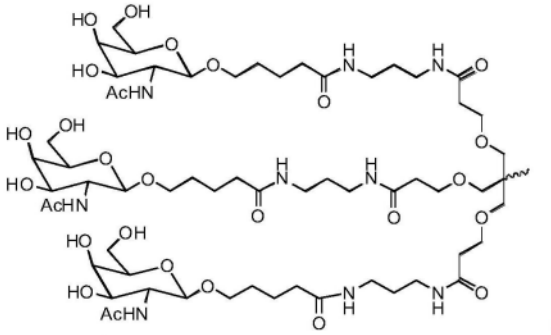
。

[0615] 在一些实施方案中,RNAi剂缀合至如表11中定义且如下所示的L96:

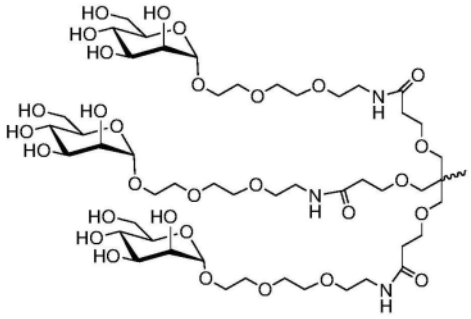
[0616] 

。

[0617] 在某些实施方案中,用于本发明的组合物及方法中的碳水化合物缀合物是选自下列所组成的组:

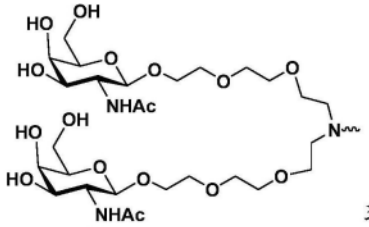


式 II;

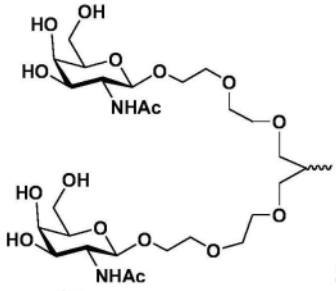


式 III;

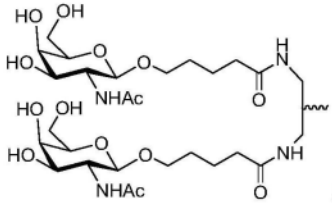
[0618]



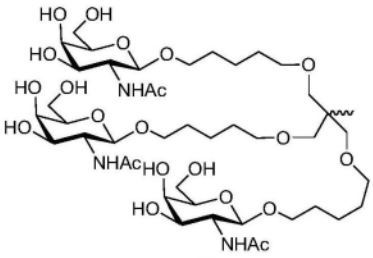
式 IV;



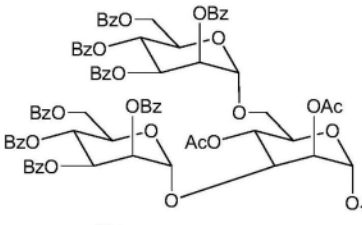
式 V;



式 VI;

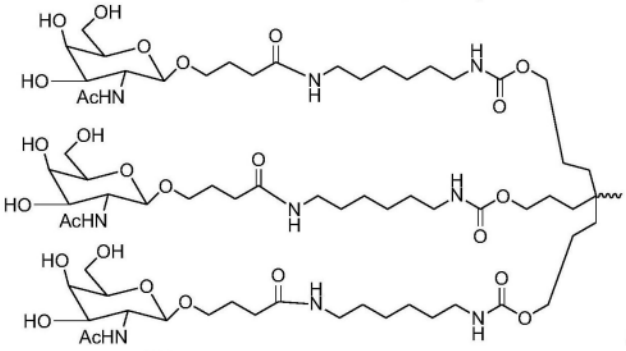


式 VII;

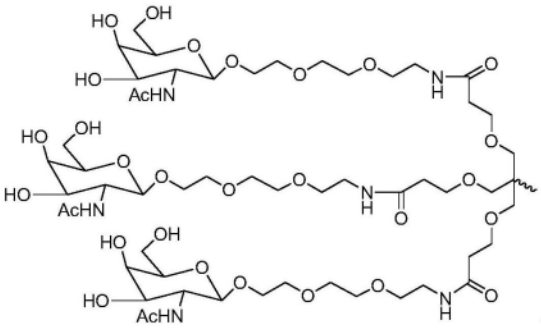


式 VIII;

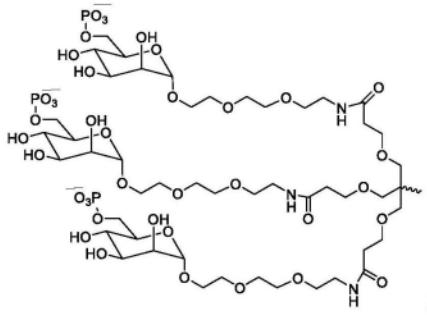
[0619]



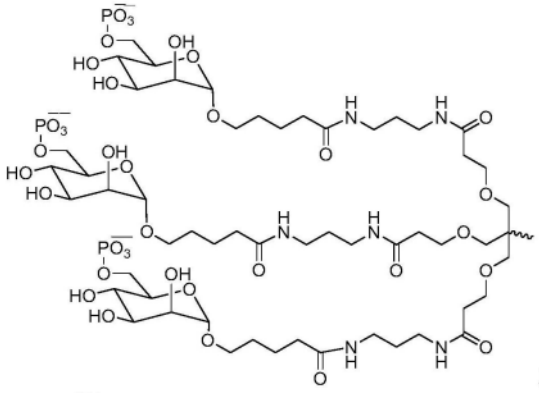
式 IX;



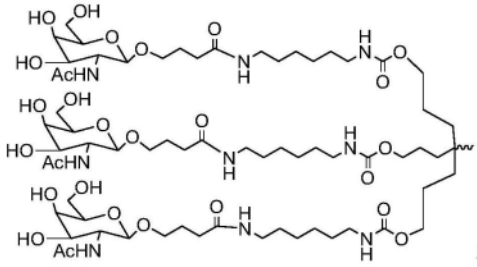
式 X;



式 XI;

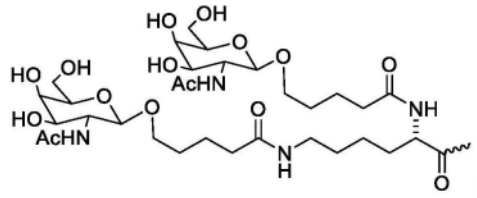


式 XII;

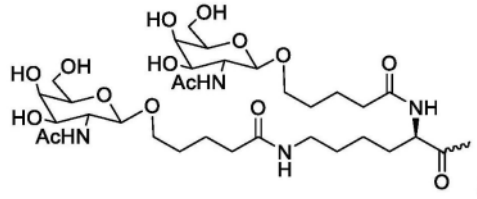


式 XIII;

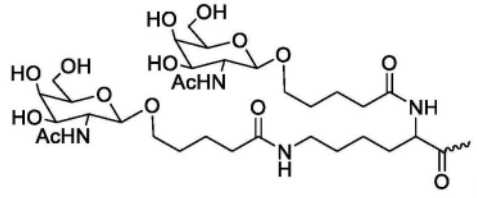
[0620]



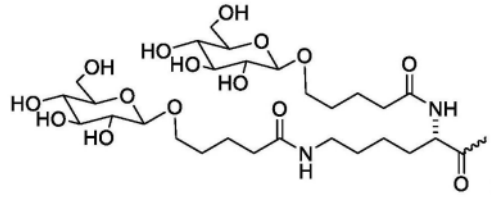
式 XIV;



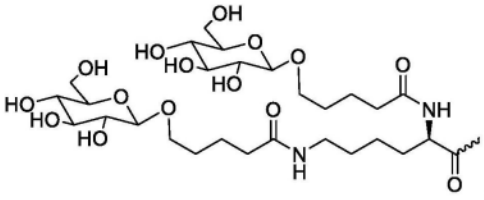
式 XV;



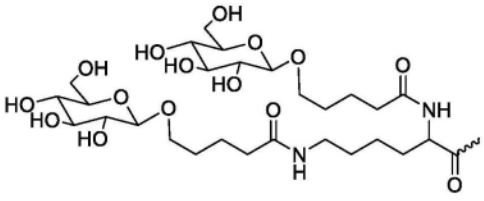
式 XVI;



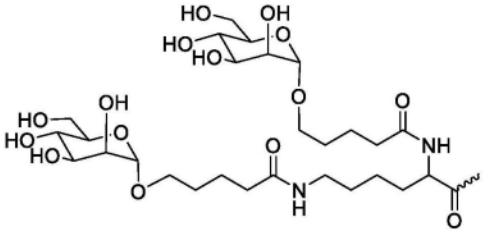
式 XVII;



式 XVIII;

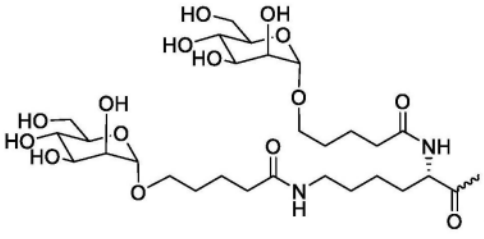


式 XIX;

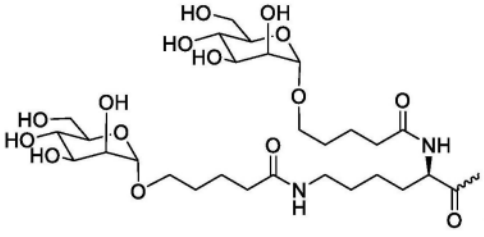


式 XX;

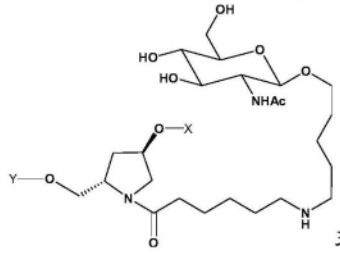
[0621]



式 XXI;

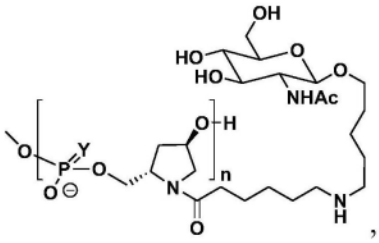


式 XXII;

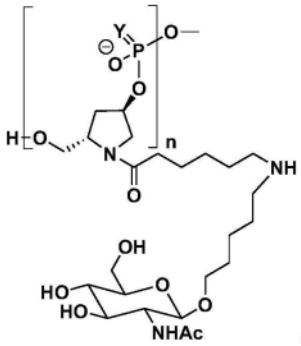


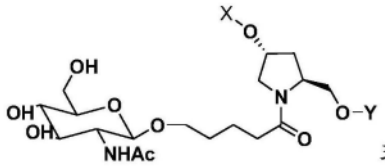
式 XXIII;

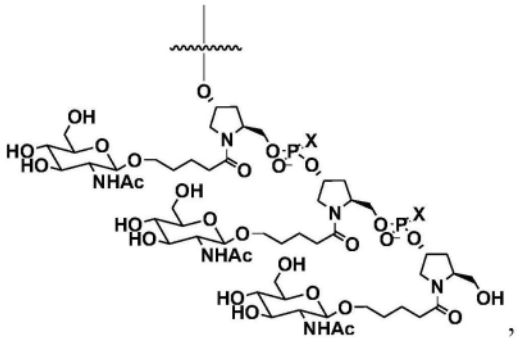
[0622]

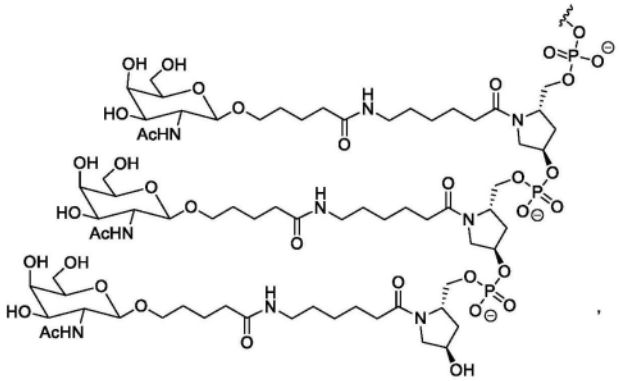


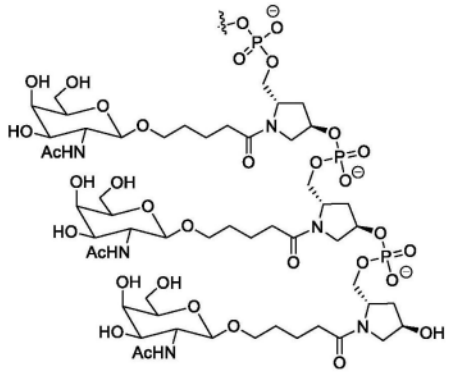
其中Y是O或S，且n是3至6(式XXIV)；

[0623]  其中Y是O或S，且n是3至6(式XXV)；

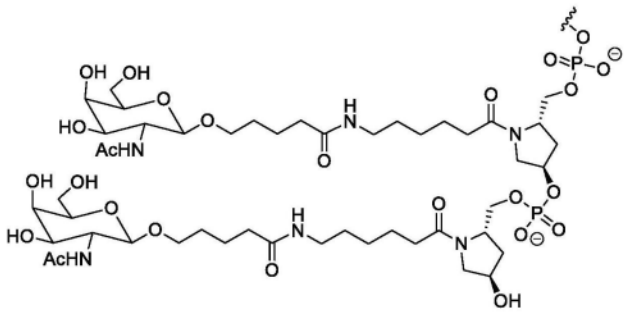
[0624]  式 XXVI;

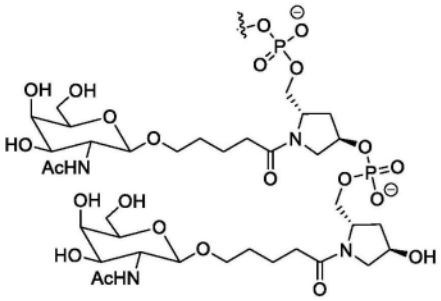
[0625]  其中X是O或S(式XXVII)；

[0626] 



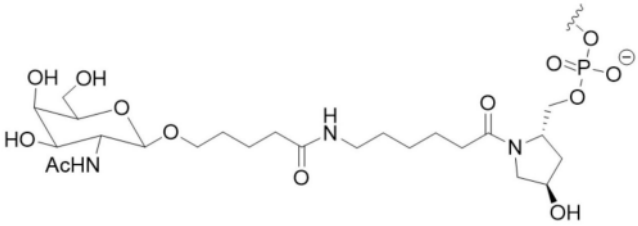
式 XXVII; 式 XXIX;

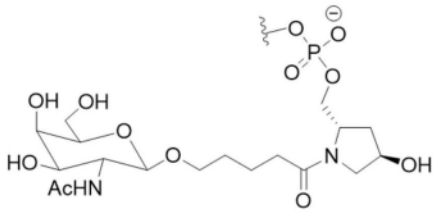




式 XXX;

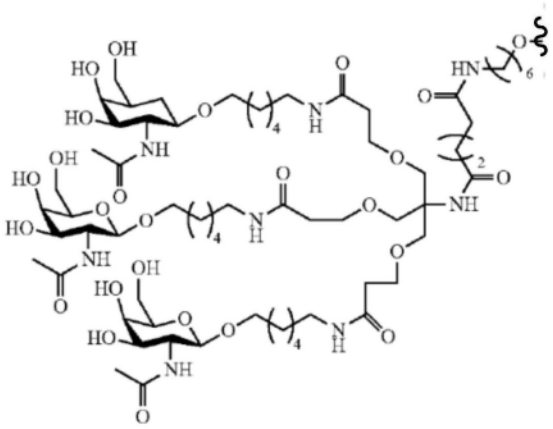
式 XXXI;

[0627]  及



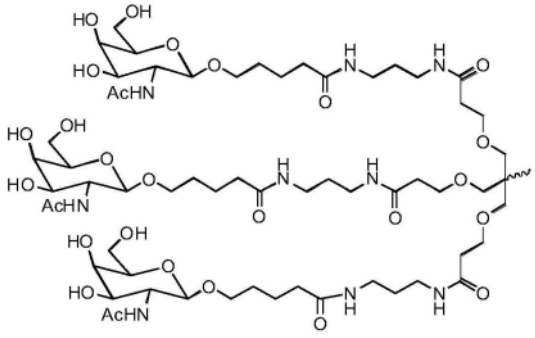
式 XXXII;

式 XXXIII;



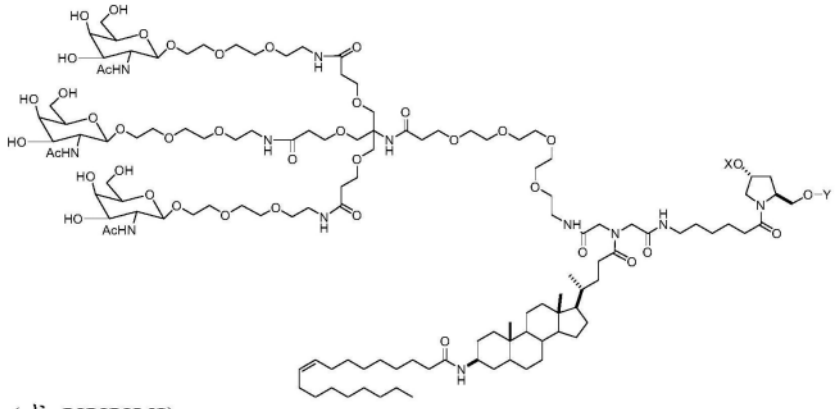
式 XXXIV。

[0628] 在某些实施方案中，用于本发明的组合物及方法中的碳水化合物缀合物是单糖。在某些实施方案中，该单糖是N-乙酰基半乳糖胺，诸如

[0629] 

式 II。

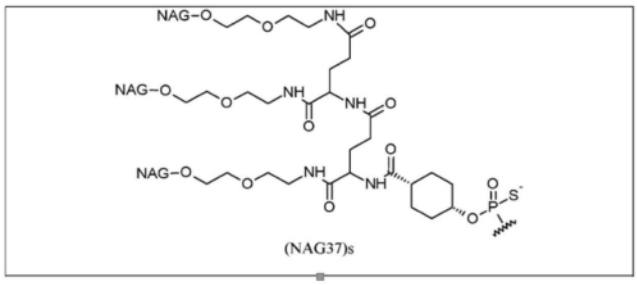
[0630] 用于本文所述的实施方案中的另一代表性碳水化合物缀合物包括但不限于，

[0631] 

(式 XXXVI)，

[0632] 当X或Y中的一个是寡核苷酸时，另一个是氢。

[0633] 在一些实施方案中，合适的配体是WO2017/0340661中披露的配体，该专利的全部内容是通过援引而并入本文。在一个实施方案中，配体包含以下结构：

[0634] 

[0635] 在某些实施方案中，本公开的RNAi剂可包括GalNAc配体，即便此类GalNAc配体目前被预测为对于本公开的优选鞘内/CNA递送途径的价值有限。

[0636] 在本发明的某些实施方案中，GalNAc或GalNAc衍生物经由单价接头附接至本发明的iRNA剂。在一些实施方案中，GalNAc或GalNAc衍生物经由二价接头附接至本发明的iRNA剂。在本发明的又一些实施方案中，GalNAc或GalNAc衍生物经由三价接头附接至本发明的iRNA剂。

[0637] 在一个实施方案中，本发明的双链RNAi剂包含一个或多个附接至该iRNA剂的GalNAc或GalNAc衍生物。GalNAc可经由接头附接至正义链或反义链上的任何核苷酸。GalNAc可附接至正义链的5’端、正义链的3’端、反义链的5’端、或反义链的3’端。在一个实施方案中，GalNAc经由三价接头附接至正义链的3’末端。

[0638] 在其他实施方案中，本发明的双链RNAi剂含多(例如，2、3、4、5或6)个GalNAc或GalNAc衍生物，其通过多个接头例如单价接头而各自独立附接至该双链RNAi剂的多个核苷酸。

[0639] 在一些实施方案中，例如，当本发明的iRNA剂的两条链都是一个更大分子的一部分且通过位于一条链的3’末端与另一条链的5’端之间的不间断核苷酸链连接而形成包含多个未配对核苷酸的发夹环时，该发夹环内的每一个未配对核苷酸可独立包含经由单价接头附接的GalNAc或GalNAc衍生物。

[0640] 在一些实施方案中，碳水化合物缀合物进一步包含如上文所述的一个或多个另外配体，例如但不限于，PK调节剂或细胞渗透肽。

[0641] 适用于本发明的额外的碳水化合物复合物(及接头)包括那些于WO 2014/179620及WO 2014/179627中所描述的，其各自的全部内容通过援引而并入本文。

[0642] D.接头

[0643] 在一些实施方案中，本文中所述的缀合物或配体可使用多种接头附接至iRNA寡核苷酸，该接头可以是可裂解的或不可裂解的。

[0644] 术语“接头”或“连接基团”意指将化合物的两个部分连接在一起如将化合物的两个部分共价附接的有机部分。接头典型地包含直接键或原子如氧或硫，单元如NR8、C(O)、C(O)NH、SO、SO_2、SO_2NH或原子的链，例如但不限于，经取代或未经取代的烷基、经取代或未经取代的烯基、经取代或未经取代的炔基、芳基烷基、芳基烯基、芳基炔基、杂芳基烷基、杂芳基烯基、杂芳基炔基、杂环基烷基、杂环基烯基、杂环基炔基、芳基、杂芳基、杂环基、环烷基、环烯基、烷基芳基烷基、烷基芳基烯基、烷基芳基炔基、烯基芳基烷基、烯基芳基烯基、烯基芳基炔基、炔基芳基烷基、炔基芳基烯基、炔基芳基炔基、烷基杂芳基烷基、烷基杂芳基烯

基、烷基杂芳基炔基、烯基杂芳基烷基、烯基杂芳基烯基、烯基杂芳基炔基、炔基杂芳基烷基、炔基杂芳基烯基、炔基杂芳基炔基、烷基杂环基烷基、烷基杂环基烯基、烷基杂环基炔基、烯基杂环基烷基、烯基杂环基烯基、烯基杂环基炔基、炔基杂环基烷基、炔基杂环基烯基、炔基杂环基炔基、烷基芳基、烯基芳基、炔基芳基、烷基杂芳基、烯基杂芳基、炔基杂芳基，其一个或多个亚甲基可由O、S、S(O)、SO_2、N(R8)、C(O)、经取代或未经取代的芳基、经取代或未经取代的杂芳基、经取代或未经取代的杂环基中断或终止；其中R8是氢、酰基、脂族或经取代的脂族。在某些实施方案中，接头是约1至24个原子、2至24个原子、3至24个原子、4至24个原子、5至24个原子、6至24个原子、6至18个原子、7至18个原子、8至18个原子、7至17个原子、8至17个原子、6至16个原子、7至16个原子、或8至16个原子。

[0645] 可裂解的连接基团在细胞外足够稳定，但当进入靶标细胞时裂解以释放被该接头保持在一起的两个部分。在优选的实施方案中，该可裂解的连接基团在靶标细胞内或在第一参考条件(其可以是例如经选择以模拟或呈现细胞内的条件)下的裂解比在受试者血液内或在第二参考条件(其可以是例如经选择以模拟或呈现见于血液或血清中的条件)下的裂解快至少约10倍、20倍、30倍、40倍、50倍、60倍、70倍、80倍、90倍或更高、或至少约100倍。

[0646] 可裂解的接头基团是对于裂解剂例如pH、氧化还原电位或可降解分子的存在敏感者的。通常，与在血清或血液中相比，裂解剂在细胞内更为普遍或以更高水平或活性被发现。此类降解剂的示例包括：氧化还原剂，其是选择用于特定的底物或其不具有底物特异性，包括例如细胞内存在的通过还原来降解氧化还原可裂解连接基团的氧化酶、还原酶或还原剂如硫醇；酯酶；内切酶，或可创建酸性环境的剂如导致pH为5或更低的那些；可通过作为通用酸、肽酶(其可以是底物特异性的)及磷酸酶起作用而水解或降解酸可裂解的连接基团的酶。

[0647] 可裂解的连接基团如二硫键可能对于pH敏感。人血清的pH为7.4，而细胞内平均pH略低，为7.1至7.3的范围。内体具有酸性更强的pH，为5.5至6.0的范围；而溶酶体具有酸性甚至更强的pH，约为5.0。一些接头将具有可裂解的连接基团，该连接基团在优选的pH裂解，从而在细胞内将阳离子脂质从该配体释放出来或将该阳离子脂质释放到所希望的细胞区室内。

[0648] 接头可包括可通过特定酶裂解的可裂解的连接基团。并入接头的可裂解的连接基团的类型可取决于待作为靶标的细胞。例如，肝脏靶向配体可通过包括酯基的接头而连接至阳离子脂质。肝细胞富含酯酶，因此接头在肝细胞中将比在不富含酯酶的细胞类型内更有效地被裂解。其他富含酯酶的细胞类型包括肺、肾皮质及睾丸的细胞。

[0649] 当靶向细胞类型是富含肽酶的如肝细胞及滑膜细胞时，可使用含有肽键的接头。

[0650] 通常，可通过测试降解剂(或条件)裂解候选连接基团的能力而评估该候选可裂解连接基团的适用性。还希望地是也测试该候选可裂解连接基团在血液中或当与其他非靶标组织接触时的抵抗裂解的能力。因此，可测定第一条件与第二条件间的裂解相对敏感性，其中该第一条件被选择为靶标细胞内的裂解标志物，且该第二条件被选择为其他组织或生物流体例如血液或血清中的裂解标志物。这些评估可在无细胞的系统内、细胞内、细胞培养物内、器官或组织培养物内、或在整个动物体内进行。可能有用者是，在无细胞或培养条件下进行初始评估，并通过在整体动物体内的进一步评估而证实。在优选的实施方案中，可用的候选化合物在细胞内(或在选择以模拟细胞内条件的体外条件下)的裂解比在血液或血清

中(或在选择以模拟细胞外条件的体外条件下)快至少约2倍、4倍、10倍、20倍、30倍、40倍、50倍、60倍、70倍、80倍、90倍或约100倍。

[0651] i.氧化还原可裂解的连接基团

[0652] 在某些实施方案中,可裂解的连接基团是氧化还原可裂解的连接基团,其当还原或氧化时被裂解。可经还原裂解的连接基团的示例是二硫连接基团(-S-S-)。为了确定候选可裂解连接基团是否是合适的"可还原裂解的连接基团"或例如是否适用于与特定iRNA部分及特定靶向剂合用,可查看本文中所述的方法。例如,可通过以二硫苏糖醇(DTT)或使用该领域中已知试剂的还原剂温育来评估候选者,该温育模拟在细胞例如靶标细胞内将会观察到的裂解速率。这些候选物还可在经选择以模拟血液或血清条件的条件下评估。在一个实施方案中,候选化合物在血液中被裂解至多约10%。在其他实施方案中,可用的候选化合物在细胞内(或在选择以模拟细胞内条件的体外条件下)的降解比在血液或血清中(或在选择以模拟细胞外条件的体外条件下)快至少约2倍、4倍、10倍、20倍、30倍、40倍、50倍、60倍、70倍、80倍、90倍或约100倍。候选化合物的裂解速率可使用标准酶动力学分析在经选择以模拟细胞内介质的条件下测定,并将其与在经选择以模拟细胞外介质的条件下所测的进行比较。

[0653] ii.基于磷酸酯的可裂解连接基团

[0654] 在某些实施方案中,可裂解的接头包含基于磷酸酯的可裂解连接基团。基于磷酸酯的可裂解连接基团通过降解或水解该磷酸酯基团的剂而裂解。在细胞内裂解磷酸酯基团的剂的示例是酶如细胞内的磷酸酶。磷酸酯基连接基团的示例是-O-P(O)(ORk)-O-、-O-P(S)(ORk)-O-、-O-P(S)(SRk)-O-、-S-P(O)(ORk)-O-、-O-P(O)(ORk)-S-、-S-P(O)(ORk)-S-、-O-P(S)(ORk)-S-、-S-P(S)(ORk)-O-、-O-P(O)(Rk)-O-、-O-P(S)(Rk)-O-、-S-P(O)(Rk)-O-、-S-P(S)(Rk)-O-、-S-P(O)(Rk)-S-、-O-P(S)(Rk)-S-。优选的实施方案是-O-P(O)(OH)-O-、-O-P(S)(OH)-O-、-O-P(S)(SH)-O-、-S-P(O)(OH)-O-、-O-P(O)(OH)-S-、-S-P(O)(OH)-S-、-O-P(S)(OH)-S-、-S-P(S)(OH)-O-、-O-P(O)(H)-O-、-O-P(S)(H)-O-、-S-P(O)(H)-O-、-S-P(S)(H)-O-、-S-P(O)(H)-S-、-O-P(S)(H)-S-。优选的实施方案是-O-P(O)(OH)-O-。这些候选物可使用与上述的那些类似的方法评估。

[0655] iii.酸可裂解的连接基团

[0656] 在某些实施方案中,可裂解的接头包含酸可裂解的连接基团。酸可裂解的连接基团是在酸性条件下被裂解的连接基团。在优选的实施方案中,酸可裂解的连接基团是在pH为约6.5或更低(例如,约6.0、5.75、5.5、5.25、5.0或更低)的酸性环境中被裂解,或被剂诸如可用作通用酸的酶裂解。在细胞中,特定的低pH细胞器如内体及溶酶体,可提供用于酸可裂解的连接基团的裂解环境。酸可裂解的连接基团的示例包括但不限于腙类、酯类、及氨基酸的酯类。酸可裂解的连接基团可具有通式-C=NN-、C(O)O、或-OC(O)。优选的实施方案是,当附接至酯的氧(烷氧基)的碳是芳基时,经取代的烷基、或叔烷基如二甲基戊基或叔丁基。这些候选物可使用与上述的那些类似的方法评估。

[0657] iv.基于酯的可裂解连接基团

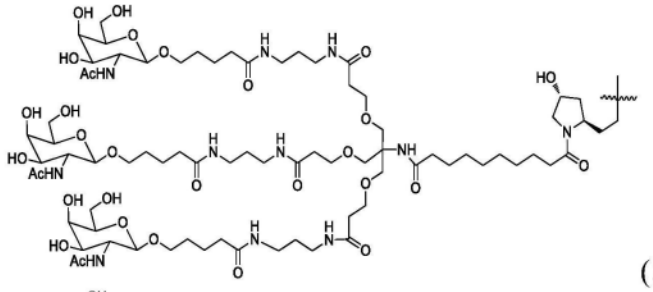
[0658] 在某些实施方案中,可裂解的接头包含基于酯的可裂解连接基团。基于酯的可裂解连接基团由酶诸如酯酶或酰胺酶在细胞内裂解。基于酯的可裂解连接基团的示例包括但不限于亚烷基、亚烯基和亚炔基的酯类。酯可裂解的连接基团可具有通式-C(O)O-或-OC

(O)-。这些候选物可使用与上述的那些类似的方法评估。

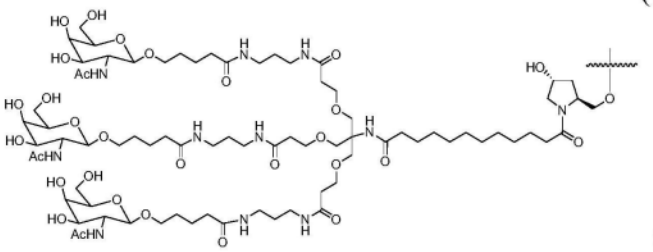
[0659] v.基于肽的可裂解连接基团

[0660] 在又一实施方案中,可裂解的接头包含基于肽的可裂解连接基团。基于肽的可裂解连接基团由酶诸如肽酶及蛋白酶在细胞内裂解。基于肽的可裂解基团是在氨基酸间形成以获得寡肽(例如,二肽、三肽等)及多肽的肽键。基于肽的可裂解基团不包括酰胺基团(-C(O)NH-)。酰胺基团可在任意亚烷基、亚烯基或亚炔基之间形成。肽键在氨基酸间形成以获得肽及蛋白质的特定类型的酰胺键。基于肽的裂解基团通常限定为在氨基酸间形成而获得肽及蛋白质的肽键(即,酰胺键),且不包括该完整酰胺官能基团。基于肽的可裂解连接基团具有通式-NHCHRAC(O)NHCHRBC(O)-(SEQ ID NO:__),其中RA与RB是两个相邻氨基酸的R基团。这些候选物可使用与上述的那些类似的方法评估。

[0661] 在一些实施方案中,本发明的iRNA通过接头缀合至碳水化合物。本发明的组合物及方法的具有接头的iRNA碳水化合物缀合物的非限制性示例包括,但不限于,

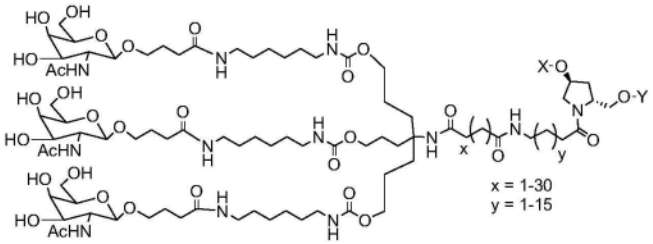


(式 XXXVII)、

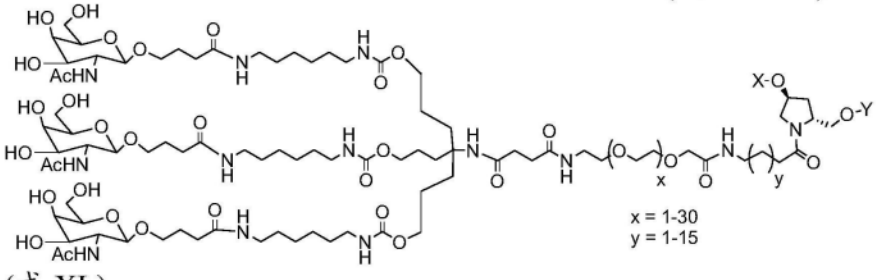


(式 XXXVIII)、

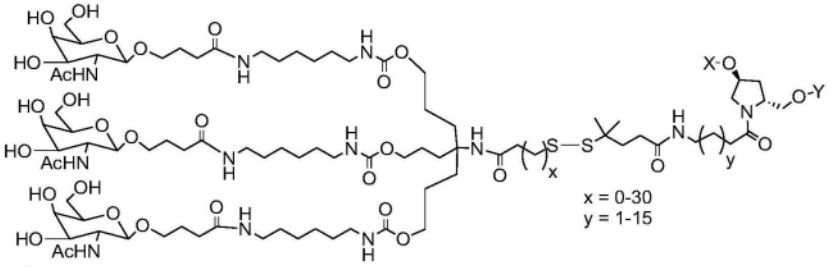
[0662]



(式 XXXIX)、

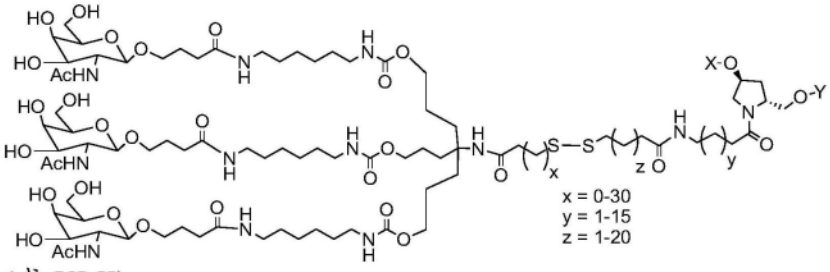


(式 XL)、

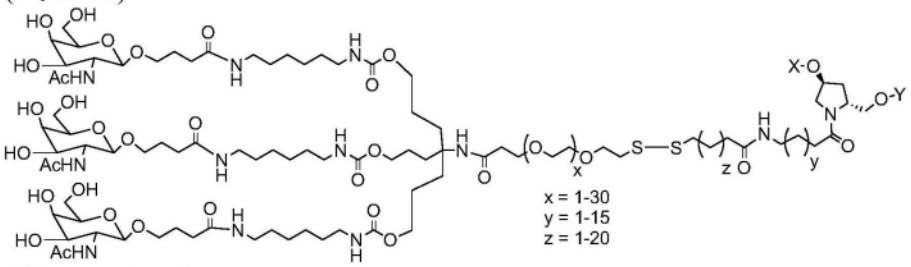


(式 XLI)、

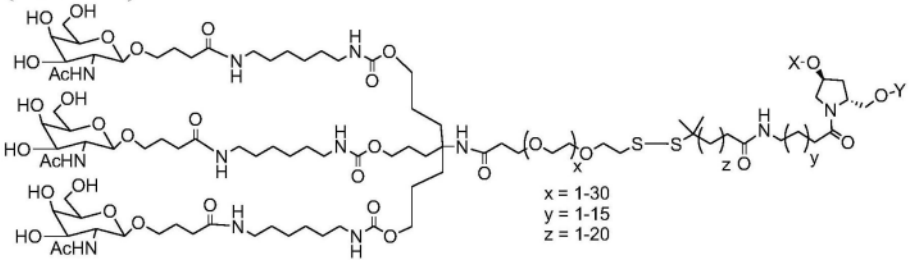
[0663]



(式 XLII)、



(式 XLIII)、及

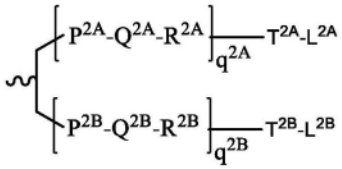


[0664] (式 XLIV),其中,X或Y中的一个是寡核苷酸,且另一个是氢。

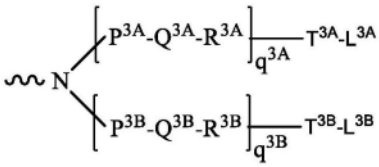
[0665] 在本发明的组合物及方法的某些实施方案中,配体是一种或多种通过二价或三价分支的接头附接的GalNAc(N-乙酰基半乳糖胺)衍生物。

[0666] 在某些实施方案中,本发明的dsRNA缀合至选自式(XLV)至(XLVI)中任一者所示结构组成的组的二价或三价分支的接头:

[0667] 式 XXXXV 式 XLVI

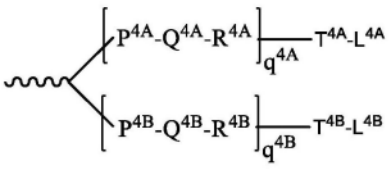


IV)

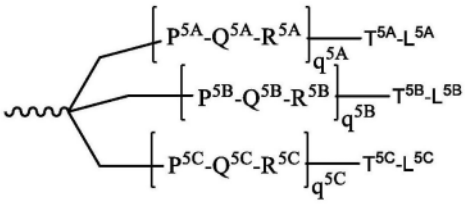


(V)

[0668]



式 XLVII



式 XLVIII

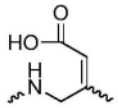
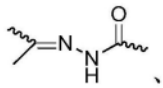
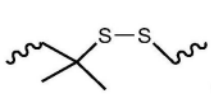
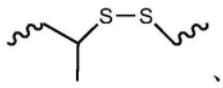
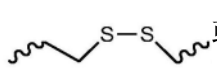
[0669] 其中：

[0670] q2A、q2B、q3A、q3B、q4A、q4B、q5A、q5B及q5C于每次出现时独立表示0至20，其中，该重复单元可以是相同或不同；

[0671] P^{2A}、P^{2B}、P^{3A}、P^{3B}、P^{4A}、P^{4B}、P^{5A}、P^{5B}、P^{5C}、T^{2A}、T^{2B}、T^{3A}、T^{3B}、T^{4A}、T^{4B}、T^{4A}、T^{5B}、T^{5C}于每次出现时独立为不存在、CO、NH、O、S、OC(O)、NHC(O)、CH_2、CH_2NH或CH_2O；

[0672] Q^{2A}、Q^{2B}、Q^{3A}、Q^{3B}、Q^{4A}、Q^{4B}、Q^{5A}、Q^{5B}、Q^{5C}于每次出现时是不存在、亚烷基、经取代的亚烷基(其中，一个或多个亚甲基可通过O、S、S(O)、SO_2、$N(R^N)$、C(R')=C(R")、C≡C或C(O)的一个或多个中断或封端)；

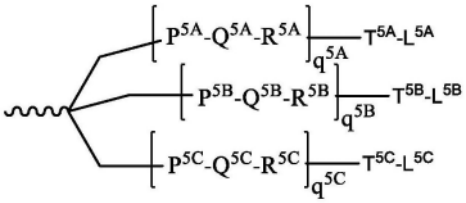
[0673] R^{2A}、R^{2B}、R^{3A}、R^{3B}、R^{4A}、R^{4B}、R^{5A}、R^{5B}、R^{5C}于每次出现时是不存在、NH、O、S、CH_2、C(O)O、C(O)NH、$NHCH(R^a)C(O)$、-C(O)-$CH(R^a)$-NH-、CO、CH=N-O、


、

、

、

、

或杂环基；

[0674] L^{2A}、L^{2B}、L^{3A}、L^{3B}、L^{4A}、L^{4B}、L^{5A}、L^{5B}及L^{5C}表示配体；即，于每次出现时各自独立为单糖(诸如GalNAc)、二糖、三糖、四糖、寡糖、或多糖；且R^a是H或氨基酸侧链。三价缀合GalNAc衍生物尤其可与RNAi剂合用，以用于抑制靶标基因的表现，例如式(XLIX)的那些：

[0675] 式XLIX

[0676]



[0677] 式，

[0678] 其中，L^{5A}、L^{5B}及L^{5C}表示单糖，诸如GalNAc衍生物。

[0679] 合适的缀合GalNAc衍生物的二价及三价分支的接头的示例包括但不限于，上文作为式II、VII、XI、X、及XIII而列举的结构。

[0680] 教导RNA缀合物的制备的代表性美国专利包括但不限于，美国专利第4,828,979号、第4,948,882号、第5,218,105号、第5,525,465号、第5,541,313号、第5,545,730号、第5,

552,538号、第5,578,717号、第5,580,731号、第5,591,584号、第5,109,124号、第5,118,802号、第5,138,045号、第5,414,077号、第5,486,603号、第5,512,439号、第5,578,718号、第5,608,046号、第4,587,044号、第4,605,735号、第4,667,025号、第4,762,779号、第4,789,737号、第4,824,941号、第4,835,263号、第4,876,335号、第4,904,582号、第4,958,013号、第5,082,830号、第5,112,963号、第5,214,136号、第5,082,830号、第5,112,963号、第5,214,136号、第5,245,022号、第5,254,469号、第5,258,506号、第5,262,536号、第5,272,250号、第5,292,873号、第5,317,098号、第5,371,241号、第5,391,723号、第5,416,203号、第5,451,463号、第5,510,475号、第5,512,667号、第5,514,785号、第5,565,552号、第5,567,810号、第5,574,142号、第5,585,481号、第5,587,371号、第5,595,726号、第5,597,696号、第5,599,923号、第5,599,928号、第5,688,941号、第6,294,664号、第6,320,017号、第6,576,752号、第6,783,931号、第6,900,297号、第7,037,646号、第8,106,022号，其各自的全部内容通过援引而并入本文。

[0681] 给定化合物的所有位置不是必需经一致修饰的，且事实上，超过一种前述修饰可并入单个化合物中或甚至并入iRNA的单个核苷处。本发明还包括作为嵌合化合物的iRNA化合物。

[0682] 在本发明的语境中，“嵌合”iRNA化合物或“嵌合体”是iRNA化合物，优选是dsRNA剂，其含有两个或更多个化学上截然不同的区域，各自由至少一个单体单元构成(在dsRNA化合物的情形中，该单体单元是核苷酸)。这些iRNA典型含有至少一个区域，其中，该RNA经修饰以赋予该iRNA以增加的对核酸酶降解的抗性、增加的细胞摄取、或增加的与靶标核酸的亲和性。iRNA的附加区域可用作能裂解RNA:DNA杂交体或RNA:RNA杂交体的酶的底物。例如，RNase H是细胞的核酸内切酶，其裂解RNA:DNA双链体的RNA链。因此，RNase H的激活导致RNA靶标的裂解，从而极大地提升对基因表达的iRNA抑制的效率。因此，当使用嵌合dsRNA时，使用较短的iRNA往往可获得与使用杂交至相同靶标区域的硫代磷酸酯脱氧dsRNA相当的结果。RNA靶标的裂解可通过凝胶电泳常规检测，且若需要，可将凝胶电泳与该领域中已知的相关核酸杂交技术合用。

[0683] 在某些例子中，iRNA的RNA可通过非配体基团予以修饰。大量非配体分子已经缀合至iRNA以提升iRNA的活性、细胞分布或细胞摄取，且执行此类缀合的过程可在科技文献中获得。此类非配体部分已经包括脂质部分，如胆固醇(Kubo,T.et al.,Biochem.Biophys.Res.Comm.,2007,365(1):54-61;Letsinger et al.,Proc.Natl.Acad.Sci.USA,1989,86:6553)、胆酸(Manoharan et al.,Bioorg.Med.Chem.Lett.,1994,4:1053)、硫醚例如己基-S-三苯甲基硫醇(Manoharan et al.,Ann.N.Y.Acad.Sci.,1992,660:306;Manoharan et al.,Bioorg.Med.Chem.Let.,1993,3:2765)、硫代胆固醇(Oberhauser et al.,Nucl.Acids Res.,1992,20:533)、脂肪链例如十二烷二醇或十一烷基残基(Saison-Behmoaras et al.,EMBO J,1991,10:1111;Kabanov et al.,FEBS Lett.,1990,259:327;Svinarchuk et al.,Biochimie,1993,75:49-54)、磷脂质例如二-十六烷基-rac-甘油或1,2-二-O-十六烷基-rac-甘油-3-磷酸三乙铵(Manoharan et al.,Tetrahedron Lett.,1995,36:3651;Shea et al.,Nucl.Acids Res.,1990,18:3777)、聚胺或聚乙二醇链(Manoharan et al.,Nucleosides&Nucleotides,1995,14:969)、或金刚烷乙酸(Manoharan et al.,Tetrahedron Lett.,1995,36:3651)、棕

棕酰基部分(Mishra et al.,Biochim.Biophys.Acta,1995,1264:229)、或十八烷基胺或己基氨基-羰氧基胆固醇部分(Crooke et al.,J.Pharmacol.Exp.Ther.,1996,277:923)。教导此类RNA缀合物的制备的代表性美国专利列述于上文中。典型的缀合策略涉及在该序列的一个或多个位置携带氨基接头的RNA的合成。氨基随后与使用合适的偶联剂或活化剂缀合的分子反应。缀合反应可使用仍键合至固体支持物的RNA执行,或在RNA于溶液相中裂解后执行。通过HPLC进行的RNA缀合物的纯化典型提供纯缀合物。

[0684] V.本公开的RNAi剂的递送

[0685] 本公开的RNAi剂至细胞(例如受试者例如人类受试者(例如,有此需要的受试者,诸如具有HTT相关疾患例如亨廷顿病的受试者)体内的细胞)的递送可通过大量不同路径实现。例如,可通过将细胞与本公开的RNAi剂在体外或体内接触而执行递送。体内递送也可通过将包含RNAi剂例如dsRNA的组合物给药至受试者而直接执行。或者,体内递送可通过将编码并引导该RNAi剂的表达的一种或多种载体给药而间接执行。这些选择在下文中进一步讨论。

[0686] 通常,递送核酸分子的任意方法(体外或体内)可适用于与本公开的RNAi剂合用(参见,例如,Akhtar S.and Julian RL.,(1992)Trends Cell.Biol.2(5):139-144及WO94/02595,其通过援引而以其整体并入本文)。对于体内递送,为了递送RNAi剂而考虑的因素包括,例如,所递送的剂的生物学稳定性、非特异性效果的预防、及所递送的剂于靶标组织内的累积。RNAi剂的非特异性效果可通过局部给药而最小化,例如,通过直接注射或移植入组织内或外用给药该制剂。局部给药至治疗位点将该剂的局部浓度最大化,限制该剂与可能受该剂伤害或可降解该剂的系统性组织接触,且允许以较低的剂量给药RNAi剂。若干研究已经显示,局部给药RNAi剂实现了成功的基因敲低的产物。例如,在食蟹猕猴体内通过眼内注射(Tolentino,MJ.et al.,(2004)Retina 24:132-138)及小鼠体内通过玻璃体内注射(Reich,SJ.et al.(2003)Mol.Vis.9:210-216)进行的VEGF dsRNA的眼内递送,两者都显示防止年龄相关性黄斑变性实验模型中的新血管生成。此外,在小鼠体内进行dsRNA的直接肿瘤内注射减小肿瘤体积(Pille,J.et al.(2005)Mol.Ther.11:267-274)且可延长荷瘤小鼠的生存期(Kim,WJ.et al.,(2006)Mol.Ther.14:343-350;Li,S.et al.,(2007)Mol.Ther.15:515-523)。RNA干扰也已经显示通过直接注射进行的局部递送的成功(Dorn,G.et al.,(2004)Nucleic Acids 32:e49;Tan,PH.et al.(2005)Gene Ther.12:59-66;Makimura,H.et a.l(2002)BMC Neurosci.3:18;Shishkina,GT.,et al.(2004)Neuroscience 129:521-528;Thakker,ER.,et al.(2004)Proc.Natl.Acad.Sci.U.S.A.101:17270-17275;Akaneya,Y.,et al.(2005)J.Neurophysiol.93:594-602)以及通过鼻内给药而成功递送至肺部(Howard,KA.et al.,(2006)Mol.Ther.14:476-484;Zhang,X.et al.,(2004)J.Biol.Chem.279:10677-10684;Bitko,V.et al.,(2005)Nat.Med.11:50-55)。对于系统性给药RNAi剂用于治疗疾病,该RNA可经修饰或者使用药物递送系统递送;两种方法都起作用以防止dsRNA被内核酸酶及外核酸在体内快速降解。RNA的修饰或药物载剂也可容许RNAi剂对靶组织的靶向效应并避免不希望的脱靶效应(例如,不希望受缚于理论,如本文所述的GNA的用途已经被鉴定为去稳定化dsRNA的种子区域,导致此类dsRNA相对于脱靶效应的上靶效应的倾向增强,因为此类脱靶效应被该种子区域去稳定化作用显著弱化)。RNAi剂可通过化学缀合至亲脂性基团诸如

胆固醇而被修饰,以增强细胞摄取并且防止降解。例如,将被缀合至亲脂性胆固醇部分的对抗ApoB的RNAi剂系统性注射至小鼠体内,导致肝脏及空肠两处的apoB mRNA的敲低(Soutschek,J.et al.,(2004)Nature 432:173-178)。已经显示,将RNAi剂缀合至适配体能抑制前列腺癌模型小鼠体内的肿瘤生长并介导肿瘤衰退(McNamara,JO.et al.,(2006)Nat.Biotechnol.24:1005-1015)。在作为另一种选择的实施方案中,RNAi剂可使用药物递送系统如纳米颗粒、树枝状聚合物、聚合物、脂质体、或阳离子递送系统进行递送。带正电的阳离子递送系统促进分子RNAi剂(带负电)的结合,还增强在带负电的细胞膜处的相互作用,以允许该细胞对RNAi剂的有效摄取。阳离子脂质、树枝状聚合物或聚合物可键合至RNAi剂,或经引入以形成封装iRNA的囊泡或微团(参见,例如,Kim SH.et al.,(2008)Journal of Controlled Release 129(2):107-116)。囊泡或微团的形成进一步防止当系统性给药时RNAi剂的降解。制作及给药阳离子-RNAi剂复合物的方法完全处于本领域技术人员的能力范围内(参见,例如,Sorensen,DR.,et al.(2003)J.Mol.Biol 327:761-766;Verma,UN.et al.,(2003)Clin.Cancer Res.9:1291-1300;Arnold,ASet al.(2007)J.Hypertens.25:197-205,其均通过援引而整体并入本文)。可用于RNAi剂的全身递送的药物递送系统的一些非限制性示例包括DOTAP(Sorensen,DR.,et al(2003),如上;Verma,UN.et al.,(2003),如上)、Oligofectamine“固体核酸脂质颗粒”(Zimmermann,TS.et al.,(2006)Nature 441:111-114)、心磷脂(Chien,PY.et al.,(2005)Cancer Gene Ther.12:321-328;Pal,A.et al.,(2005)Int J.Oncol.26:1087-1091)、聚乙烯亚胺(Bonnet ME.et al.,(2008)Pharm.Res.Aug 16Epub ahead of print;Aigner,A.(2006)J.Biomed.Biotechnol.71659)、Arg-Gly-Asp(RGD)肽(Liu,S.(2006)Mol.Pharm.3:472-487)、及聚酰胺基胺(Tomalia,DA.et al.,(2007)Biochem.Soc.Trans.35:61-67;Yoo,H.et al.,(1999)Pharm.Res.16:1799-1804)。在一些实施方案中,RNAi剂与环糊精形成用于系统性给药的复合物。给药的方法及RNAi剂与环糊精的药物组合物可见于美国专利第7,427,605号,其通过援引而以其整体并入本文。

[0687] 本公开的某些方面涉及降低细胞中C9orf72靶标基因的表达的方法,包括使该细胞与本公开的双链RNAi剂接触。在一个实施方案中,细胞是肝外细胞,任选是CNS细胞。

[0688] 本公开的另一方面涉及降低受试者体内C9orf72靶标基因的表达的方法,包括向该受试者给药本公开的双链RNAi剂。

[0689] 本公开的另一方面涉及治疗患有CNS疾患的受试者的方法,包括向该受试者给药治疗有效量的本公开的双链C9orf72靶向RNAi剂,从而治疗受试者。可通过本公开的方法治疗的示例性CNS疾患包括C9orf72相关疾病。

[0690] 在一个实施方案中,双链RNAi剂是经鞘内给药。通过双链RNAi剂的鞘内给药,该方法降低脑(例如,纹状体)或脊髓组织例如皮质、小脑、颈部脊髓、腰部脊髓及胸部脊髓中的C9orf72靶标基因表达。

[0691] 为了便于阐述,本章节中的制剂、组合物及方法主要讨论经修饰的siRNA化合物。然而,可以理解,这些制剂、组合物及方法可使用其他siRNA化合物例如未修饰的siRNA化合物实践,并且此实践处于本公开范围内。可以通过多种途径将包括RNAi剂的组合物递送至受试者。示例性途径包括:鞘内、静脉内、外用、直肠内、经肛门、阴道内、鼻内、肺部内及眼内。

[0692] 本公开的RNAi剂可并入适用于给药的药物组合物中。此类组合物典型包括一种或多种RNAi剂及药学可接受的载剂。如本文所用,短语“药学可接受的载剂”旨在包括与药物给药相容的任意及全部溶剂、分散介质、涂层、抗细菌剂及抗真菌剂、等张剂及吸收延迟剂等。此类用于药学活性物质的介质及剂的使用是该领域中熟知的。除非任意常规介质或剂与活性化合物不兼容,否则预期其在组合物中的使用。补充活性化合物也可并入组合物中。

[0693] 本公开的药物组合物可经由多种途径给药,取决于局部治疗或系统性治疗是否是所希望的,且取决于待治疗的面积。给药可以是外用(包括眼部给药、阴道给药、直肠给药、鼻腔内给药、透皮给药)、口服或肠胃外给药。肠胃外给药包括静脉滴注、皮下注射、腹膜内注射或肌肉内注射,或者鞘内给药或脑室内给药。

[0694] 可选择给药的途径及位点以增强靶向性。例如,为了靶向肌肉细胞,肌肉内注射至感兴趣的肌肉中将是合乎逻辑的选择。通过给药喷雾剂形式的RNAi剂,可靶向肺细胞。通过以RNAi剂涂覆球囊导管以及机械地引入RNA,可靶向血管内皮细胞。

[0695] 用于外用给药的制剂可包括透皮贴剂、软膏剂、洗剂、乳霜剂、凝胶剂、滴剂、栓剂、喷雾剂、液体及粉末剂。常规药物载剂、水性基质、粉末基质、油状基质、增稠剂等可以是必需的或想要的。经涂覆的保险套(condom)、手套等也可以是有用者。

[0696] 用于口服给药的组合物包括粉末剂或颗粒剂、水中的悬浮剂或溶液剂、糖浆剂、酏剂或非水性介质、片剂、胶囊剂、菱锭剂或锭剂。在片剂的情况下,可使用的载剂包括乳糖、柠檬酸钠及磷酸的盐。各种崩解剂诸如淀粉及润滑剂诸如硬脂酸镁、月桂基硫酸钠及滑石常用于片剂中。对于胶囊形式的口服给药,可用的稀释剂是乳糖及高分子量聚乙二醇。当需要使用水性悬浮剂进行口服使用时,可将核酸组合物与乳化剂及悬浮剂组合。若需要,可添加某些甜味剂或芳香剂。

[0697] 用于鞘内或脑室内给药的组合物可包括无菌水溶液,该无菌水溶液也可含有缓冲剂、稀释剂及其他合适的添加剂。

[0698] 用于肠胃外给药的制剂可包括无菌水溶液,该无菌水溶液也可含有缓冲剂、稀释剂及其他合适的添加剂。脑室内注射可通过例如附接至储器的脑室内导管得以促进。对于静脉内用途,可控制溶质的总浓度以使得制备物等张。

[0699] 在一个实施方案中,siRNA化合物例如双链siRNA化合物或ssiRNA化合物、组合物的给药是肠胃外给药例如静脉内给药(例如,作为推注或作为扩散性输注)、皮内给药、腹膜内给药、肌肉内给药、鞘内给药、脑室内给药、颅内给药、皮下给药、跨黏膜给药、口含给药、舌下给药、内窥镜给药、直肠给药、口服给药、阴道给药、外用给药、肺部给药、鼻腔内给药、尿道给药或眼部给药。给药可由实施者提供或由另一个人例如医疗从业人员提供。药品可以可测量的剂量提供或提供在递送定量剂量的分配器中。下文更详细地讨论所选择的递送模式。

[0700] A.鞘内给药

[0701] 在一个实施方案中,通过鞘内注射(即,注射至脑及脊髓组织浸没于其中的脊髓液中)递送双链RNAi剂。将RNAi剂鞘内注射至脊髓液中,可作为推注或经由迷你泵(minipump)执行,该迷你泵可植入皮肤下,提供规则且恒定的siRNA至脊髓液中的递送。脊髓液自脉络丛(其在该处产生)环绕脊髓及背根神经节下行,然后上行经过小脑及皮质到达蛛网膜粒,在该处可离开CNS,从而完成循环,依据所注射化合物的尺寸、稳定性及溶解度,经鞘内递送

的分子可击中整个CNS范围内的靶标。

[0702] 在一些实施方案中，鞘内腔给药是经由泵进行。泵可以是经外科手术植入的渗透泵。在一个实施方案中，将渗透泵植入脊髓管的蛛网膜下腔内以促进鞘内给药。

[0703] 在一些实施方案中，鞘内给药是经由药用的鞘内递送系统进行，该系统包括含有一定体积药剂的储器，以及配置为递送该储器中所含药剂的一部分的泵。关于这一鞘内递送系统的更多细节可见于WO 2015/116658中，该专利通过援引以其整体并入本文。

[0704] 对于不同的靶标基因，经鞘内注射的RNAi剂的量可变，并且必须施加的适宜量可能必须关于每种靶标基因单独确定。典型地，这一量是10μg至2mg，优选50μg至1500μg，更优选100μg至1000μg的范围。

[0705] B.编码本公开的RNAi剂的载体

[0706] 靶向C9orf72基因的RNAi剂可从插入DNA或RNA载体的转录单元表达(参见，例如，Couture,A,et al.,TIG.(1996),12:5-10;WO 00/22113、WO 00/22114及US 6,054,299)。表达优选是持续(数月或更久)，取决于所使用的特定构建体及靶标组织或细胞类型。这些转基因可作为线性构建体、环状质体、或病毒载体而引入，其可以是整合载体或非整合载体。该转基因还可构建为允许其被作为线粒体外质体而被遗传(Gassmann,et al.,(1995) Proc.Natl.Acad.Sci.USA 92:1292)。

[0707] RNAi剂的个体链或多链可从表达载体的启动子转录。若两个分开的链待表达以生成例如dsRNA，则可将两个分开的表达载体共同引入(例如，通过转染或感染共同引入)靶标细胞内。或者，dsRNA的每一个体链可通过位于相同表达质体上的两种启动子转录。在一个实施方案中，dsRNA表达为反向重复聚核苷酸，其通过接头聚核苷酸序列缀合，使得该dsRNA具有茎环结构。

[0708] RNAi剂表达载体通常是DNA质体或病毒载体。与真核细胞兼容的表达载体，优选是与脊椎动物细胞兼容的那些，可用来生产用于表达本文所述的RNAi剂的重组构建体。RNAi剂表达载体的递送可以是系统性的，如通过静脉内或肌肉内给药；通过给药至从该患者外植的靶标细胞，之后重新引入患者体内；或通过任何其他容许引入所希望的靶标细胞内的手段。

[0709] 可与本文所述方法及组合物合用的病毒载体系统是包括但不限于，(a)腺病毒载体；(b)逆转录病毒载体，包括但不限于慢病毒载体、莫罗尼鼠(moloney murine)白血病病毒等；(c)腺相关病毒载体；(d)单纯疱疹病毒载体；(e)SV40载体；(f)多瘤病毒载体；(g)乳头状瘤病毒载体；(h)小核糖核酸病毒载体；(i)痘病毒载体，如天花如牛痘病毒载体，或禽痘如金丝雀痘或鸡痘病毒载体；以及(j)帮手依赖性或裸线病毒载体。复制缺陷病毒也可以是有利的。不同的载体将变为或不变为并入该细胞的基因组内。若必要，这些构建体可包括用于转染的病毒序列。或者，该构建体可并入能进行附加型复制(episomal replication)的载体如EPV载体及EBV载体内。用于RNAi剂的重组表达的构建体通常将会需要调节元件例如启动子、增强子等，以确保该RNAi剂在靶标细胞内的表达。对于载体及构建体所考虑的其他方面是该领域中已知的。

[0710] VI.本发明的药物组合物

[0711] 本公开还包括，包括本公开的RNAi剂的药物组合物及制剂。在一个实施方案中，本文提供含有如本文所述的RNAi剂以及药学可接受的载剂的药物组合物。含有RNAi剂的药物

组合物可用于治疗与C9orf72的表达或活性相关的疾病或疾患，例如，C9orf72相关疾病。

[0712] 在一些实施方案中，本发明的药物组合物是无菌的。在另一实施方案中，本发明的药物组合物是无热原的。

[0713] 这些药物组合物是基于递送模式而配制。一个示例是将组合物配制为用于经由肠胃外递送的系统性给药，例如，通过静脉内(IV)递送、肌肉内(IM)递送或用于皮下(subQ)递送。另一示例是配制为用于直接递送至CNS中的组合物，例如通过鞘内或玻璃体内注射途径，任选通过输注入脑(例如，纹状体)内，诸如通过连续泵输注。

[0714] 本公开的药物组合物可以足以抑制C9orf72基因表达的剂量给药。通常，本公开的RNAi剂的适当剂量将是约0.001至约200.0毫克每千克接受者体重每天的范围内，通常是约1至50毫克每千克体重每天的范围内。

[0715] 重复剂量方案可包括规则地给药治疗量的RNAi剂，如每个月一次至每六个月一次。在某些实施方案中，RNAi剂是约每季给药一次(即，约每三个月给药一次)至约每年给药两次。

[0716] 在初始的治疗方案(例如，负荷剂量)后，治疗的实施频次可降低。

[0717] 在其他实施方案中，单一剂量的药物组合物可以是长期持续的，使得后续剂量是以不超过1、2、3或4个月或更久的间隔给药。在本公开的一些实施方案中，单一剂量的本公开的药物组合物是每个月给药一次。在本公开的其他实施方案中，单一剂量的本公开的药物组合物是每个季度给药一次至每年给药两次。

[0718] 本领域普通技术人员应知悉，某些因素可影响有效治疗受试者所需的剂量及时机，这些因素包括但不限于疾病或病变的严重性、先前的治疗、受试者的一般健康情况/或年龄、以及存在的其他疾病。此外，使用治疗有效量的组合物治疗受试者可包括单一治疗或一系列治疗。

[0719] 在小鼠遗传学中的进展已经生成大量用于研究多种将会受益于C9orf72表达的降低的人类疾病诸如ALS及FTD的小鼠模型。此类模型可用于RNAi剂的体内测试，以及用于确定治疗有效剂量。合适的啮齿动物模型是本领域中已知的，并且包括例如那些描述于例如Cepeda,et al.(ASN Neuro(2010)2(2):e00033)及Pouladi,et al.(Nat Reviews(2013)14:708)中的。

[0720] 本公开的药物组合物可经由多种途径给药，取决于局部治疗或系统性治疗是否为所希望的，且取决于待治疗的面积。给药可以是局部给药(例如，通过透皮贴剂)；肺部给药，例如通过粉末或气溶胶的吸入或吹入，包括通过喷雾器；气管内给药、鼻内给药、表皮给药及透皮给药；口服或肠道外给药。肠胃外给药包括静脉内、动脉内、皮下、腹膜内或肌肉内注射或输注；真皮下给药，例如经由植入装置；或颅内给药，例如通过脑实质内给药、鞘内腔给药或脑室内给药。

[0721] 可以靶向具体组织诸如CNS(例如，脑的神经元、神经胶质或血管组织)的模式递送RNAi剂。

[0722] 用于外用给药的药物组合物及制剂可包括透皮贴剂、软膏剂、洗剂、乳霜剂、凝胶剂、滴剂、栓剂、喷雾剂、液体及粉末剂。常规药物载剂、水性基质、粉末基质、油状基质、增稠剂等可以是必需的或所希望的。经涂覆的保险套、手套等也可以是有用的。适宜的外用制剂包括下述的那些，其中本公开所提出的RNAi剂与外用递送剂如脂质、脂质体、脂肪酸、脂肪

酸酯、类固醇、螯合剂及表面活性剂混合。适宜的脂质及脂质体包括中性的(如,二油酰基磷脂DOPE乙醇胺、二肉豆蔻酰基卵磷脂DMPC、二硬脂酰基卵磷脂)、阴性的(如,二肉豆蔻酰基磷脂甘油DMPG)及阳离子性的(如,二油酰基四甲基氨基丙基DOTAP及二油酰基磷脂乙醇胺DOTMA)。本公开提出的RNAi剂可封装在脂质体内,或可与脂质体尤其是阳离子脂质体形成复合物。另选地,RNAi剂可与脂质尤其是阳离子脂质复合。适宜的脂肪酸及酯类包括但不限于花生四烯酸、油酸、花生酸、月桂酸、辛酸、癸酸、肉豆蔻酸、棕榈酸、硬脂酸、亚油酸、苏子油酸、二癸酸酯、三癸酸酯、单油酸甘油酯、二月桂酸甘油酯、甘油1-单癸酸酯、1-十二烷基氮杂环庚-2-酮、酰基肉碱、酰基胆碱、或其C_{1-20}烷基酯(如,肉豆蔻酸异丙酯IPM)、单甘油酯、二甘油酯或药学可接受的盐。外用制剂详细描述于US 6,747,014中,其通过援引而并入本文。

[0723] A.含膜分子组件的RNAi剂制剂

[0724] 用于本公开的组合物及方法中的RNAi剂可配制为用于在膜分子组件如脂质体或微团中递送。如本文中所用,术语"脂质体"是指由排列为至少一个双层例如一个双层或多个双侧的两亲性脂质构成的囊泡。脂质体包括单层或多层囊泡,其具有从亲脂性材料形成的膜及水性内腔。水性部分含有RNAi剂组合物。亲脂性材料将水性内腔与水性外部分离,而该水性外部并不包括RNAi剂组合物,但在一些示例中,可包括RNAi剂组合物。脂质体是有用于将活性成分转移并递送至起作用位点。因为脂质体膜在结构上类似于生物膜,当将脂质体施用至组织时,脂质体双层与细胞膜的双层融合。随着脂质体与细胞的融合的进行,包括RNAi剂的内部水性内容物被递送至细胞内,在该处,RNAi剂可特异性地结合靶标RNA并可介导RNAi。在一些情况下,脂质体也被特异性地靶向例如以将RNAi剂引导至特定细胞类型。

[0725] 含有RNAi剂的脂质体可通过多种方法制备。在一个示例中,将脂质体的脂质成分溶解在洗涤剂中,使得与该脂质成分形成微团。例如,该脂质成分可以是两亲性阳离子脂质或脂质复合物。该洗涤剂可具有高临界微团浓度且可以是非离子性的。示例性的洗涤剂包括胆酸盐、CHAPS、辛基葡萄糖苷、脱氧胆酸盐、及月桂酰肌氨酸。随后将RNAi剂制剂加入包括该脂质成分的微团中。该脂质上的阳离子性基团与RNAi剂相互作用,并缩合在RNAi剂周围以形成脂质体。缩合之后,例如通过渗析移除该洗涤剂,以得到RNAi剂的脂质体性制剂。

[0726] 若必要,可在缩合反应过程中,例如通过受控添加而加入有助于缩合的载剂化合物。例如,该载剂化合物可以是除核酸的外的聚合物(如,精胺或精三胺)。还可调节pH以辅助缩合。

[0727] 生产稳定的聚核苷酸递送囊泡的方法(该方法将聚核苷酸/阳离子脂质复合物作为递送囊泡的结构性成分而并入)进一步描述于例如WO 96/37194中,其全部内容通过援引而并入本文。脂质体制剂还可包括下列中所述的例示性方法的一个或多个方面:Felgner,P.L.et al.,(1987)Proc.Natl.Acad.Sci.USA 8:7413-7417;美国专利第4,897,355号;美国专利第5,171,678号;Bangham et al.,(1965)M.Mol.Biol.23:238;Olson et al.,(1979)Biochim.Biophys.Acta 557:9;Szoka et al.,(1978)Proc.Natl.Acad.Sci.75:4194;Mayhew et al.,(1984)Biochim.Biophys.Acta 775:169;Kim et al.,(1983)Biochim.Biophys.Acta 728:339;以及Fukunaga et al.,(1984)Endocrinol.115:757。常用的制备其尺寸适合用作递送囊泡的脂质聚集体的技术包括超音波处理及冻融加挤出(参见,例如,Mayeret al.,(1986)Biochim.Biophys.Acta 858:161)。当一致性地小(50至

200nm)且相对均匀的聚集体是所希望的时,可使用微流体化(Mayhewet al.,(1984)Biochim.Biophys.Acta775:169)。这些方法可容易地适用于将RNAi剂制剂封装入脂质体中。

[0728] 脂质体落入两个大类中。阳离子脂质体是带正电的脂质体,其与带负电的核酸分子相互作用以形成稳定的复合物。带正电的核酸/脂质体复合物结合至带负电的细胞表面,且在内体中被内化。由于内体中的酸性pH,该脂质体破裂,将其内容物释放到细胞质中(Wang et al.(1987)Biochem.Biophys.Res.Commun.,147:980-985)。

[0729] pH敏感或带负电的脂质体包埋核酸而非与核酸形成复合物。由于核酸及脂质两者均带相似的电荷,出现排斥而非形成复合物。尽管如此,一些核酸仍被包埋至这些脂质体的水性内腔中。pH敏感的脂质体已经用来将编码胸苷激酶基因的核酸递送至培养物中的细胞单层。在靶标细胞中检测到外源性基因的表达(Zhou et al.(1992)Journal of Controlled Release,19:269-274)。

[0730] 一种主要类型的脂质体性组合物包括除天然衍生的卵磷脂外的磷脂质。例如,中性脂质体组合物可由二肉豆蔻酰基卵磷脂(DMPC)或二棕榈酰基卵磷脂(DPPC)形成。阴离子性脂质体组合物通常由二肉豆蔻酰基磷脂酰甘油形成,而阴离子性促融合脂质体主要由二油酰基磷脂酰乙醇胺(DOPE)形成。另一类型的脂质体性组合物由卵磷脂(PC)诸如,例如,大豆PC及蛋PC形成。另一类型由磷脂质或卵磷脂或胆固醇的混合物形成。

[0731] 在体外及体内将脂质体引入细胞内的其他方法的示例是包括美国专利第5,283,185号;美国专利第5,171,678号;WO 94/00569;WO 93/24640;WO 91/16024;Felgner,(1994)J.Biol.Chem.269:2550;Nabel,(1993)Proc.Natl.Acad.Sci.90:11307;Nabel,(1992)Human Gene Ther.3:649;Gershon,(1993)Biochem.32:7143;以及Strauss,(1992)EMBO J.11:417。

[0732] 也已经检查非离子性脂质体系统,尤其是包含非离子性表面活性剂及胆固醇的系统,以确定它们在将药物递送至皮肤中的用途。包含Novasome™ I(甘油二月桂酸酯/胆固醇/聚氧乙烯-10-硬脂基醚)及Novasome™II(甘油二硬脂酸酯/胆固醇/聚氧乙烯-10-硬脂基醚)的非离子性脂质体制剂是用来将环孢素-A递送至小鼠皮肤的真皮内。结果表明,此类非离子性脂质体系统是有效促进环孢素-A沉积在皮肤的不同层内(Hu et al.,(1994)S.T.P.Pharma.Sci.,4(6):466)。

[0733] 脂质体还包括“立体稳定化的”脂质体,如本文中使用,该术语指代包含一种或多种空间化脂质的脂质体,当该空间化脂质被并入脂质体中时,其导致循环寿命比缺失此类空间化脂质的脂质体提升。立体稳定化的脂质体的示例是下述的那些:其中,脂质体的形成囊泡的脂质部分中的一部分(A)是包含一种或多种糖脂质,如单唾液酸神经节苷脂G_{M1},或(B)是使用一种或多种亲水性聚合物如聚乙二醇(PEG)部分予以衍生。尽管不希望受缚于任何特定理论,在该领域中知晓,至少对于含有神经节苷脂、鞘磷脂、或PEG衍生的脂质的立体稳定化的脂质体,这些立体稳定化的脂质体的提升的循环半衰期是源于网状内皮系统(RES)的细胞对其的摄取降低(Allen et al.,(1987)FEBS Letters,223:42;Wu et al.,(1993)Cancer Research,53:3765)。

[0734] 多种包含一种或多种磷脂质的脂质体是该领域中已知的。Papahadjopoulos et al.(Ann.N.Y.Acad.Sci.,(1987),507,64)报导了单唾液酸神经节苷脂G_{M1}、硫酸半乳糖脑苷

脂及磷脂酰肌醇改善脂质体的血液半衰期的能力。这些发现是通过下列阐述:Gabizon et al.(Proc.Natl.Acad.Sci.U.S.A.,(1988),85,:6949)。美国专利第4,837,028号及WO 88/04924,两者均授予Allen等人,其披露包含(1)鞘磷脂及(2)神经节苷脂G_{M1}或硫酸半乳糖脑苷脂的脂质体。美国专利第5,543,152号(Webb et al.)披露包含鞘磷脂的脂质体。包含1,2-sn-二肉豆蔻酰基卵磷脂的脂质体是披露于WO 97/13499(Lim et al)中。

[0735] 在一个实施方案中,使用阳离子性脂质体。阳离子脂质体具备能融合至细胞膜的优点。尽管非阳离子脂质体不能有效地与质膜融合,但其可在体内被巨噬细胞摄取,且可用以将RNAi剂递送至巨噬细胞。

[0736] 脂质体的进一步的优点包括:从天然磷脂质获得的脂质体是生物相容且生物可降解者;脂质体可合并多种水及脂溶性药物;脂质体可保护封装在其内部区室中的RNAi剂不被代谢及降解(Rosoff,in"Pharmaceutical Dosage Forms",Lieberman,Rieger and Banker(Eds.),1988,volume 1,p.245)。在脂质体制剂的制备中的重要考虑是脂质表面电荷、囊泡尺寸、及脂质体的水性体积。

[0737] 带正电的合成阳离子脂质,氯化N-[1-(2,3-二油酰基氧)丙基]-N,N,N-三甲基铵(DOTMA),可用以形成小脂质体,该小脂质体自发与核酸反应以形成脂质-核酸复合物,该复合物能与组织培养细胞的细胞膜的带负电的脂质融合,从而完成RNAi剂的递送(参见,例如,Felgner,P.L.et al.,(1987)Proc.Natl.Acad.Sci.USA 8:7413-7417及美国专利第4,897,355号的对DOTMA及其与DNA合用的说明)。

[0738] 一种DOTMA类似物,1,2-双(油酰基氧基)-3-(三甲基氨)丙烷(DOTAP),可与磷脂质合用以形成复合有DNA的囊泡。Lipofectin™(Bethesda Research Laboratories,Gaithersburg,Md.)是用于将高度阴离子性核酸递送至活体组织培养细胞内的有效剂,该细胞包含带正电的DOTMA脂质体,而该脂质体自发与带负电的聚核苷酸相互作用以形成复合物。当使用带足够多正电的脂质体时,所得复合物上的静电荷也为正。以此途径制备的带正电的复合物自发地附接至带负电的细胞表面,与质膜融合,且有效地将官能性核酸递送至例如组织培养细胞内。另一可商购的阳离子脂质,1,2-双(油酰基氧基)-3,3-(三甲基氨)丙烷("DOTAP")(Boehringer Mannheim,Indianapolis,Indiana)与DOTMA的不同之处在于其油酰基部分是通过酯而非醚连接者。

[0739] 其他经报导的阳离子脂质化合物包括已经缀合至多种部分的那些,包括,例如,已经缀合至两种类型的脂质之一的羧基精胺,且包括化合物如5-羧基精胺基甘氨酸十八油酰基酰胺("DOGS")(Transfectam™,Promega,Madison,Wisconsin)及二棕榈酰基磷脂酰乙醇胺5-羧基精胺基-酰胺("DPPES")(参见,例如,美国专利第5,171,678号)。

[0740] 另一阳离子脂质复合物是包括具胆固醇的脂质衍生物("DC-Chol"),其已经与DOPE组合而配制在脂质体内(参见,Gao,X.and Huang,L.,(1991)Biochim.Biophys.Res.Commun.179:280)。脂质聚赖氨酸,由缀合聚赖氨酸至DOPE而制成,已经被报导其在存在血清的情况下的转染中有效(Zhou,X.et al.,(1991)Biochim.Biophys.Acta 1065:8)。对于某些细胞,据称这些含有经缀合的阳离子脂质的脂质体显现较低的毒性且提供比含DOTMA的组合物更有效的转染。其他可商购的阳离子脂质产物包括DMRIE及DMRIE-HP(Vical,La Jolla,California)及Lipofectamine(DOSPA)(Life Technology,Inc.,Gaithersburg,Maryland)。其他适用于递送寡核苷酸的阳离子脂质描述

于WO 98/39359及WO 96/37194中。

[0741] 脂质体制剂尤其适用于外用给药，脂质体呈现比其他制剂优越的若干优点。此类优点包括，相对于对所给药的药物的高度系统性吸收，副作用降低；所给药的药物在所希望的靶标处的累积增加；以及，将RNAi剂给药至皮肤内的能力。在一些实施方式中，脂质体用于将RNAi剂递送至表皮细胞，且还用以提升RNAi剂至真皮组织内如皮肤内的渗透。例如，该脂质体可外用施加。已经有文献报导典型的配制为脂质体的药物至皮肤的递送(参见，例如，Weiner et al.,(1992)Journal of Drug Targeting,vol.2,405-410及du Plessis et al.,(1992)Antiviral Research,18:259-265;Mannino,R.J.and Fould-Fogerite,S.,(1998)Biotechniques6:682-690;Itani,T.et al.,(1987)Gene 56:267-276;Nicolau,C.et al.(1987)Meth.Enzymol.149:157-176;Straubinger,R.M.and Papahadjopoulos,D.(1983)Meth.Enzymol.101:512-527;Wang,C.Y.and Huang,L.,(1987)Proc.Natl.Acad.Sci.USA 84:7851-7855)。

[0742] 也已经检查非离子性脂质体系统，尤其是包含非离子性表面活性剂及胆固醇的系统，以确定它们在将药物递送至皮肤中的用途。包含Novasome I(甘油二月桂酸酯/胆固醇/聚氧乙烯-10-硬脂基醚)及Novasome II(甘油二硬脂酸酯/胆固醇/聚氧乙烯-10-硬脂基醚)的非离子性脂质体制剂用来将药物递送至小鼠皮肤的真皮内。此类具有RNAi剂的制剂可用于治疗皮肤疾患。

[0743] 包括RNAi剂的脂质体可制成可高度变形的。该变形性使脂质体能够渗透穿过小于该脂质体的平均半径的孔。例如，传递体(transfersome)是一种类型的可变形脂质体。传递体可通过将表面边缘活化剂，一般为表面活性剂，加入标准脂质体性组合物而制成。包括RNAi剂的传递体可通过例如注射而在皮下递送，以将RNAi剂递送至皮肤的角质细胞。为了横跨哺乳动物的总皮层，脂质囊泡必需在合适的透皮梯度的影响下穿透一系列微孔，每一微孔具有小于50nm的直径。此外，由于脂质的特性，这些传递体可以是自优化(调适至孔如皮肤内的孔的形状)、自修复，且可频繁到达其靶标而不片段化，且一般为自携带。

[0744] 适用于本公开的其他制剂描述于下列美国临时专利申请案中：2008年1月2日递交的第61/018,616号、2008年1月2日递交的第61/018,611号、2008年3月26日递交的第61/039,748号、2008年4月22日递交的第61/047,087号及2008年5月8日递交的第61/051,528号。2007年10月3日递交的PCT申请案第PCT/US2007/080331号也描述了适用于本公开的制剂。

[0745] 传递体是又一类型的脂质体，其是可高度变形的脂质聚集体，对于药物递送囊泡而言，其是有吸引力的候选。传递体可描述为脂质液滴，其可变形性如此高以至于它们能轻易地渗透穿过小于该液滴的孔。传递体可适应其所使用的环境，例如，它们是自优化(调适至孔如皮肤内的孔的形状)、自修复，且可频繁到达其靶标而不片段化，且一般是自携带。为了制作传递体，可将表面边缘活化剂，一般为表面活性剂，加入标准脂质体性组合物。传递体已经用以将血清白蛋白递送至皮肤。已经显示，传递体介导的血清白蛋白的递送与将含有血清白蛋白的溶液进行皮下注射同样有效。

[0746] 表面活性剂可广泛用于制剂(诸如本文所述的那些)，尤其是乳液(包括微乳液)及脂质体中。对包括天然及合成者在内的多种不同类型的表面活性剂的特性进行分类及排序的最常见途径，是通过使用亲水/亲脂平衡(HLB)进行。亲水性基团(也称为“头部”)的固有

特性是提供将制剂中所用的不同表面活性剂归类的最有用手段(Rieger,in Pharmaceutical Dosage Forms,Marcel Dekker,Inc.,New York,N.Y.,1988,p.285)。

[0747] 如果表面活性剂分子未经离子化,则其分类为非离子性表面活性剂。非离子性表面活性剂可广泛应用于医药及化妆产品中,且可在宽范围的pH下使用。通常,它们的HLB值是2至约18的范围,取决于它们的结构。非离子性表面活性剂包括非离子性酯类如乙二醇酯类、丙二醇酯类、甘油酯类、聚甘油酯类、失水山梨醇酯类、蔗糖酯类、及经乙氧基化的酯类。非离子性烷醇酰胺类及醚类如脂肪醇乙氧基化物、经丙氧基化的醇类、及经乙氧基化/丙氧基化的嵌段聚合物也包括于这一类中。聚氧乙烯表面活性剂是非离子性表面活性剂类别中最常见的成员。

[0748] 如果当将表面活性剂分子溶解或分散于水中时其携带负电荷,则该表面活性剂是分类为阴离子性。阴离子性表面活性剂包括羧酸酯类如皂类、酰基乳酸酯类、氨基酸的酰基酰胺类、硫酸的酯类如硫酸烷基酯及经乙氧基化的硫酸烷基酯、磺酸盐类如烷基苯磺酸盐类、酰基羟乙基磺酸盐类、酰基酒石酸盐类及磺基琥珀酸盐类、及磷酸盐类。阴离子性表面活性剂类别的最重要的成员是烷基硫酸盐类及皂类。

[0749] 如果当将表面活性剂分子溶解或分散于水中时其携带正电荷,则该表面活性剂是分类为阳离子性。阳离子表面活性剂包括季铵盐类及经乙氧基化的胺类。季铵盐类是本类别中最常用的成员。

[0750] 如果表面活性剂具有携带正电荷或负电荷的能力,则该表面活性剂分类为两性。两性表面活性剂包括丙烯酸衍生物、经取代的烷基酰胺类、N-烷基甜菜碱类及磷脂类。

[0751] 表面活性剂在药物产品、制剂及乳液中的使用已经得以回顾(Rieger,in Pharmaceutical Dosage Forms,Marcel Dekker,Inc.,New York,N.Y.,1988,p.285)。

[0752] 用于本公开的方法中的RNAi剂也可提供为微团制剂。本文中,“微团”定义为特定类型的分子组件,其中,两性分子排列为球状结构,使得这些分子的疏水性部分全部朝向内侧,留下亲水性部分与周围的水相接触。如果环境是疏水性,则存在逆向排列。

[0753] 适用于通过跨真皮膜递送的混合微团制剂可通过将siRNA组合物的水性溶液、碱金属的C_8至C_{22}烷基硫酸盐、及形成微团的化合物混合而制备。示例性的形成微团的化合物包括卵磷脂;透明质酸;透明质酸、乙醇酸、乳酸的药学可接受的盐类;洋甘菊提取物;黄瓜提取物;亚油酸;亚麻酸;单油酸甘油酯;单油酸酯类;单月桂酸酯类;玻璃苣油;月见草油;薄荷油;三羟基侧氧基胆烷基甘油及其药学可接受的盐类;甘油;聚甘油;赖氨酸;聚赖氨酸;三油酸甘油酯;聚氧乙烯醚类及其类似物;聚多卡醇烷基醚类及其类似物;鹅脱氧胆酸盐类;脱氧胆酸盐类;及其混合物。形成微团的化合物可在加入碱金属的烷基硫酸盐的同时或之后加入。为了提供较小尺寸的微团,可使用除剧烈混合外的基本上任何种类的混合形成混合微团。

[0754] 在一种方法中,制备含有siRNA组合物及至少一种碱金属的烷基硫酸盐的第一微团组合物。随后将该第一微团组合物与至少三种形成微团的化合物混合以形成混合微团组合物。在另一方法中,通过将siRNA组合物、碱金属的烷基硫酸盐、及至少一种形成微团的化合物混合,之后在剧烈混合下加入剩余的形成微团的化合物而制备。

[0755] 可将苯酚或间甲酚加至该混合微团组合物中,以稳定化制剂并防止细菌生长。或者,可将苯酚或间甲酚与形成微团的成分一起加入。在形成该混合微团组合物之后,还可加

入等张剂如甘油。

[0756] 对于将微团制剂作为喷雾剂递送，可将该制剂置于气溶胶分散器内，且该分散器填充有推进剂。该推进剂处于压力的下，在该分散器中为液体形式。调节各成分的比率，使得水性相与推进剂相成为一体，即，仅存在一相。如果存在两相，则在例如通过计量阀分散这些成分的一部分之前摇动该分散器。所分散剂量的药剂是由该计量阀推进为细小喷雾。

[0757] 推进剂可包括含氢的氯氟碳化合物、含氢的氟碳化合物、甲醚及乙醚。在某些实施方案中，可使用HFA 134a(1,1,1,2-四氟乙烷)。

[0758] 主要成分的具体浓度可通过相对简单的实验确定。对于通过口腔进行的吸收，一般所希望的是增加剂量，例如，增加为通过注射给药或通过胃肠道给药的剂量的至少两倍或三倍。

[0759] B.脂质颗粒

[0760] 本公开的RNAi剂例如dsRNA可完全封装在脂质制剂如LNP中，或封装在其他核酸-脂质颗粒内。

[0761] 如本文中所使用，术语“LNP”是指稳定的核酸-脂质颗粒。LNP典型含有阳离子脂质、非阳离子的、及预防该颗粒聚集的脂质(例如，PEG-脂质缀合物)。LNP极其有用于系统性应用，因为它们在静脉内(i.v.)注射后显现延长的循环寿命且在远程位点(例如，物理上与给药位点分隔的位点)累积。LNP包括“pSPLP”，其包括经封装的缩合剂-核酸复合物，如PCT公布第WO 00/03683号中所详述。本公开的颗粒典型具有约50nm至约150nm、更典型约60nm至约130nm、更典型约70nm至约110nm、最典型约70nm至约90nm的平均直径，且基本上无毒。此外，当本公开的核酸-脂质颗粒中存在核酸时，该核酸在水性溶液中对抗核酸酶的降解。核酸-脂质颗粒及它们的制备方法披露于例如美国专利第5,976,567号、第5,981,501号、第6,534,484号、第6,586,410号、第6,815,432号及美国专利公布第2010/0324120号以及WO 96/40964中。

[0762] 在一个实施方案中，脂质与药物的比率(质量/质量比率)(例如，脂质与dsRNA的比率)将在约1:1至约50:1、约1:1至约25:1、约3:1至约15:1，从约4:1至约10:1、约5:1至约9:1、或约6:1至约9:1的范围内。上文列举的范围之间的范围也视为本公开的一部分。

[0763] 某些用于递送RNAi的特定LNP制剂已经描述于该领域中，包括，例如，描述于例如WO 2008/042973中的“LNP01”制剂，该专利通过援引并入本文。

[0764] 其他示例性脂质-dsRNA制剂鉴定于下表中。

[0765] 表2.

[0766]

	可离子化/阳离子脂质	阳离子脂质/非阳离子脂质/胆固醇/PEG-脂质缀合物 脂质：siRNA 比
SNALP-1	l,2-二亚油烯氧基-N,N-二甲基氨基丙烷 (DLinDMA)	DLinDMA/DPPC/胆固醇/PEG-cDMA (57.1/7.1/34.4/1.4) 脂质:siRNA ~ 7:1
2-XTC	2,2-二亚油烯基-4- 二甲基氨基乙基-[1,3]-二氧杂环戊烷 (XTC)	XTC/DPPC/胆固醇/PEG-cDMA 57.1/7.1/34.4/1.4 脂质:siRNA ~ 7:1
LNP05	2,2-二亚油烯基-4-二甲基氨基乙基-[1,3]- 二氧杂环戊烷 (XTC)	XTC/DSPC/胆固醇/PEG-DMG 57.5/7.5/31.5/3.5 脂质:siRNA ~ 6:1
LNP06	2,2-二亚油烯基-4-二甲基氨基乙基-[1,3]- 二氧杂环戊烷 (XTC)	XTC/DSPC/胆固醇/PEG-DMG 57.5/7.5/31.5/3.5 脂质:siRNA ~ 11:1
LNP07	2,2-二亚油烯基-4-二甲基氨基乙基-[1,3]- 二氧杂环戊烷 (XTC)	XTC/DSPC/胆固醇/PEG-DMG 60/7.5/31/1.5, 脂质:siRNA ~ 6:1
LNP08	2,2-二亚油烯基-4-二甲基氨基乙基-[1,3]- 二氧杂环戊烷 (XTC)	XTC/DSPC/胆固醇/PEG-DMG 60/7.5/31/1.5, 脂质:siRNA ~ 11:1
LNP09	2,2-二亚油烯基-4-二甲基氨基乙基-[1,3]- 二氧杂环戊烷 (XTC)	XTC/DSPC/胆固醇/PEG-DMG 50/10/38.5/1.5 脂质:siRNA 10:1
LNP10	(3aR,5s,6aS)-N,N-二甲基-2,2-二((9Z,12Z)-十八-9,12-二烯基)四氢-3aH-环戊并[d][1,3]二氧杂环戊-5-胺 (ALN100)	ALN100/DSPC/胆固醇/PEG-DMG 50/10/38.5/1.5 脂质:siRNA 10:1
LNP11	(6Z,9Z,28Z,31Z)-三十七-6,9,28,31-四烯-19-基 4-(二甲基氨基)丁酸酯 (MC3)	MC-3/DSPC/胆固醇/PEG-DMG 50/10/38.5/1.5 脂质:siRNA 10:1
LNP12	1,1'-(2-(4-(2-((2-(双(2-羟基十二基)氨基)乙基)(2-羟基十二基)氨基)乙基)哌嗪-1-基)乙基氮烷二基)二十二烷-2-醇 (Tech G1)	Tech G1/DSPC/胆固醇/PEG-DMG 50/10/38.5/1.5 脂质:siRNA 10:1
LNP13	XTC	XTC/DSPC/Chol/PEG-DMG 50/10/38.5/1.5

[0767]

		脂质:siRNA: 33:1
LNP14	MC3	MC3/DSPC/Chol/PEG-DMG 40/15/40/5 脂质:siRNA: 11:1
LNP15	MC3	MC3/DSPC/Chol/PEG-DSG/GalNAc-PEG-DSG 50/10/35/4.5/0.5 脂质:siRNA: 11:1
LNP16	MC3	MC3/DSPC/Chol/PEG-DMG 50/10/38.5/1.5 脂质:siRNA: 7:1
LNP17	MC3	MC3/DSPC/Chol/PEG-DSG 50/10/38.5/1.5 脂质:siRNA: 10:1
LNP18	MC3	MC3/DSPC/Chol/PEG-DMG 50/10/38.5/1.5 脂质:siRNA: 12:1
LNP19	MC3	MC3/DSPC/Chol/PEG-DMG 50/10/35/5 脂质:siRNA: 8:1
LNP20	MC3	MC3/DSPC/Chol/PEG-DPG 50/10/38.5/1.5 脂质:siRNA: 10:1
LNP21	C12-200	C12-200/DSPC/Chol/PEG-DSG 50/10/38.5/1.5 脂质:siRNA: 7:1
LNP22	XTC	XTC/DSPC/Chol/PEG-DSG 50/10/38.5/1.5 脂质:siRNA: 10:1

[0768] DSPC:二硬脂酰磷脂酰胆碱

[0769] DPPC:二棕榈酰磷脂酰胆碱

[0770] PEG-DMG:PEG-二肉桂酰基甘油(C14-PEG,或PEG-C14)(PEG,平均分子量为2000)

[0771] PEG-DSG:PEG-二桂皮基甘油(C18-PEG,或PEG-C14)(PEG,平均分子量为2000)

[0772] PEG-cDMA:PEG-氨基甲酰基-1,2-二肉癸酰基氧基丙胺(PEG,平均分子量为2000)

[0773] 包含SNALP(1,2-二亚油烯氧基-N,N-二氨基氨基丙烷(DLinDMA))的制剂,描述于WO 2009/127060中,该专利通过援引而并入本文。

[0774] 包含XTC的制剂,描述于WO 2010/088537中,该专利的全部内容通过援引而并入本文。

[0775] 包含MC3的制剂,描述于例如美国专利公布第2010/0324120号中,该专利的全部内容通过援引而并入本文。

[0776] 包含ALNY-100的制剂,描述于WO 2010/054406中,该专利的全部内容通过援引而并入本文。

[0777] 包含C12-200的制剂，描述于WO 2010/129709中，该专利的全部内容通过援引而并入本文。

[0778] 用于口服给药的组合物及制剂包括粉末剂或颗粒剂、微粒剂、纳米颗粒剂、水中或非水性介质中的悬浮液或溶液、胶囊剂、软胶囊剂、袋剂、片剂、或小片剂。增稠剂、芳香剂、稀释剂、乳化剂、分散助剂或黏合剂可以是所希望的。在一些实施方案中，口服制剂是下述的那些，其中本公开提出的dsRNA与一种或多种渗透增强剂表面活性剂及螯合剂协同给药。合适的表面活性剂包括脂肪酸类或其酯类或盐类、胆汁酸类或其盐类。合适的胆汁酸类/胆汁酸盐类包括鹅脱氧胆酸(CDCA)及乌索脱氧鹅脱氧胆酸(UDCA)、胆酸、去氢胆酸、脱氧胆酸、糖胆酸、甘胆酸、甘氨脱氧胆酸、牛磺胆酸、牛磺脱氧胆酸、牛磺-24,25-二氢-褐霉酸钠及甘氨二氢褐霉酸钠。适宜的脂肪酸类是包括花生四烯酸、十一碳酸、油酸、月桂酸、辛酸、癸酸、肉豆蔻酸、棕榈酸、硬脂酸、亚油酸、苏子油酸、二癸酸酯、三癸酸酯、单油酸甘油酯、二月桂酸甘油酯、甘油1-单癸酸酯、1-十二烷基氮杂环庚-2-酮、酰基肉碱、酰基胆碱、或其单甘油酯、二甘油酯、或其药学可接受的盐(例如，钠盐)。在一些实施方案中，使用渗透增强剂的组合，例如，脂肪酸类/脂肪酸盐类与胆汁酸/胆汁酸盐类的组合。一种例示性的组合是月桂酸的钠盐、癸酸及UDCA。其他渗透增强剂包括聚氧乙烯-9-月桂基醚、聚氧乙烯-20-鲸蜡基醚。本公开提出的dsRNA可作为包括喷雾干燥颗粒在内的颗粒剂形式或复合形成微粒或纳米颗粒而经口递送。dsRNA复合剂包括聚氨基酸;聚亚胺;聚丙烯酸酯;聚丙烯酸烷基酯、聚氧杂环丁烷、聚氰基丙烯酸烷基酯;阳离子化的明胶、白蛋白、淀粉、丙烯酸酯、聚乙烯醇(PEG)及淀粉;聚氰基丙烯酸烷基酯;DEAE衍生的聚亚胺、支链淀粉、纤维素及淀粉。适宜的复合剂包括几丁聚糖、N-三甲基几丁聚糖、聚-L-赖氨酸、聚组氨酸、聚鸟氨酸、聚精氨酸、鱼精蛋白、聚乙烯基吡啶、聚硫代二乙基氨基甲基乙烯P(TDAE)、聚氨基苯乙烯(如，对-胺基)、聚(氰基丙烯酸甲酯)、聚(氰基丙烯酸乙酯)、聚(氰基丙烯酸丁酯)、聚(氰基丙烯酸异丁酯)、聚(氰基丙烯酸异己酯)、DEAE-丙烯酸甲酯、DEAE-丙烯酸己酯、DEAE-丙烯酰胺、DEAE-白蛋白及DEAE-聚葡萄糖、聚丙烯酸甲酯、聚丙烯酸己酯、聚(D,L-乳酸)、聚(DL-乳酸-共-乙醇酸(PLGA)、藻酸盐、及聚乙二醇(PEG)。dsRNA的口服制剂及其制备详细描述于美国专利第6,887,906号、U.S.2003/0027780及美国专利第6,747,014号中，其各自通过援引并入本文。

[0779] 用于肠胃外、脑实质内(至脑内)、鞘膜内、心室内或肝内给药的组合物及制剂可包括无菌水溶液，其也可含有缓冲剂、稀释剂及其他适宜的添加剂，例如但不限于，渗透增强剂、载剂化合物及其他药学可接受的载剂或赋形剂。

[0780] 本公开的药物组合物包括但不限于，溶液、乳液、及含脂质体的制剂。这些组合物可从多种组分生成，这些组分包括但不限于，预成形的液体、自乳化的固体及自乳化的半固体。特别优选当治疗APP相关疾病或疾患时靶向脑的制剂。

[0781] 可便利地以单位剂型存在的本公开的医药制剂，可根据医药工业中熟知的常规技术制备。此类技术包括将活性成分与药物载剂或赋形剂联合的步骤。通常，这些制剂是通过将活性成分与液体载剂或精细分切的固体载剂或两者均匀且紧密地联合而制备，随后，若必要，成形产物。

[0782] 本公开的组合物可配制为多种可能剂型的任一者，例如但不限于，片剂、胶囊剂、软胶囊剂、液体糖浆剂、软胶剂、栓剂、及灌肠剂。本公开的组合物也可配制为处于水性、非水性或混合介质中的悬浮液。水性悬浮液可进一步含有增加悬浮液黏度的物质，该物质包

括,例如,羧甲基纤维素钠、山梨醇或葡聚糖。悬浮液也可含有稳定剂。

[0783] C.其他制剂

[0784] i.乳液

[0785] 本公开的组合物可制备且配制为乳液。乳液是一种液体以直径通常超过0.1μm的液滴形式分散于另一种液体中的典型非均质系统(参见,例如,Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems,Allen,LV.,Popovich NG.,and Ansel HC.,2004,Lippincott Williams&Wilkins(8th ed.),New York,NY;Idson,Pharmaceutical Dosage Forms,Lieberman,Rieger and Banker(Eds.),1988,Marcel Dekker,Inc.,New York,N.Y.,第1卷,第199页;Rosoff,in Pharmaceutical Dosage Forms,Lieberman,Rieger and Banker(Eds.),1988,Marcel Dekker,Inc.,New York,N.Y.,第1卷,第245页;Block in Pharmaceutical Dosage Forms,Lieberman,Rieger and Banker(Eds.),1988,Marcel Dekker,Inc.,New York,N.Y.,第2卷,第335页;Higuchi et al.,in Remington's Pharmaceutical Sciences,Mack Publishing Co.,Easton,Pa.,1985,第301页)。乳液通常包含彼此紧密混合及分散的两个不互混液体相的双相系统。通常,乳液可以是油包水(w/o)类或水包油(o/w)类。当水性相经精细切分为小液滴并分散于整块油性相中时,所得组合物称为油包水(w/o)乳液。或者,当油性相经精细切分为小液滴并分散于整块水性相中时,所得组合物称为水包油(o/w)乳液。乳液除了含有分散相及可作为水性相、油性相存在于溶液中或本身作为单独一相的活性药物之外,还可含有额外的组分。如需要,医药赋形剂如乳化剂、稳定剂、染料及抗氧化剂也可存在于乳液中。医药乳液也可以是由超过两相构成的多乳液,例如,油包水包油(o/w/o)乳液及水包油包水(w/o/w)乳液。此类复合物配制剂往往提供简单双相乳液所不具有的某些优点。其中将o/w乳液的个体油滴容纳在小水滴内的多乳液构建w/o/w乳液。同样,油滴被容纳于稳定存在于油性连续相中的水球内的系统,提供o/w/o乳液。

[0786] 乳液的特征在于热力学稳定性小或没有。一般情况下,乳液的分散相或不连续相良好地分散在外部相或连续相中,且通过乳化剂手段或形成黏度的手段维持其形式。乳液的任一相可以是半固体或固体,如在乳液型软膏基质及乳霜剂的情形中那样。其他稳定化乳液的手段需要使用可并入乳液的任一相中的乳化剂。广义上,乳化剂可归为四类:合成表面活性剂、天然乳化剂、吸收基质、及精细分散的固体(参见,例如,Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems,Allen,LV.,Popovich NG.,and Ansel HC.,2004,Lippincott Williams&Wilkins(8th ed.),New York,NY;Idson,in Pharmaceutical Dosage Forms,Lieberman,Rieger and Banker(Eds.),1988,Marcel Dekker,Inc.,New York,N.Y.,第1卷,第199页)。

[0787] 合成表面活性剂,也称为表面活性剂,已经广泛用于乳液制剂中且已经在文献中回顾(参见,例如,Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems,Allen,LV.,Popovich NG.,and Ansel HC.,2004,Lippincott Williams&Wilkins(8th ed.),New York,NY;Rieger,in Pharmaceutical Dosage Forms,Lieberman,Rieger and Banker(Eds.),1988,Marcel Dekker,Inc.,New York,N.Y.,第1卷,第285页;Idson,in Pharmaceutical Dosage Forms,Lieberman,Rieger and Banker(Eds.),Marcel Dekker,Inc.,New York,N.Y.,1988,第1卷,第199页)。表面活性剂典型是两性的且包含亲水性部分

及疏水性部分。表面活性剂的亲水性与疏水性的比率已经定义为亲水/亲脂平衡(HLB),且是在制剂的制备中归类及选择表面活性剂的有价值的工具。基于其亲水性基团的性质,表面活性剂可归为不同的类别:非离子性、阴离子性、阳离子性及两性(参见,例如,Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems,Allen,LV.,Popovich NG.,and Ansel HC.,2004,Lippincott Williams&Wilkins(8th ed.),New York,NY;Rieger,in Pharmaceutical Dosage Forms,Lieberman,Rieger and Banker(Eds.),1988,Marcel Dekker,Inc.,New York,N.Y.,第1卷,第285页)。

[0788] 乳液制剂中使用的天然乳化剂包括羊毛脂、蜂蜡、磷脂质、卵磷脂及阿拉伯胶。吸收基质具备亲水特性,使得它们可吸收水以形成w/o乳液,而仍保持其半固体一致性,吸收基质是例如无水羊毛脂及亲水石油脂。精细切分的固体也已经用作良好的乳化剂,尤其是与表面活性剂合用或用于黏性制剂中。这些包括极性无机固体,如重金属氢氧化物、非溶胀黏土如皂土、镁铝海泡石、水辉石、高岭土、蒙脱石、胶体硅酸铝及胶体硅酸镁铝、颜料及非极性固体如碳或甘油三硬脂酸酯。

[0789] 大量非乳化材料也包括于乳液制剂中,且对乳液的特性有所贡献。这些包括脂肪、油类、蜡、脂肪酸、脂肪醇、脂肪酯、保湿剂亲水性胶体、防腐剂及抗氧化剂(Block,in Pharmaceutical Dosage Forms,Lieberman,Rieger and Banker(Eds.),1988,Marcel Dekker,Inc.,New York,N.Y.,第1卷,第335页;Idson,in Pharmaceutical Dosage Forms,Lieberman,Rieger and Banker(Eds.),1988,Marcel Dekker,Inc.,New York,N.Y.,第1卷,第199页)。

[0790] 亲水性胶体或水胶体包括天然胶及合成化合物,如多糖(例如,阿拉伯胶、琼脂、藻酸、角叉菜胶、瓜尔胶、刺梧桐胶及黄芪胶)、纤维素衍生物(例如,羧甲基纤维素及羧丙基纤维素)、及合成聚合物(例如,卡波姆、纤维素醚、及羧基乙烯基聚合物)。这些在水中分散或溶胀以形成胶体溶液,其通过形成环绕分散相液滴的强界面膜且通过增加外部相的黏度而稳定化乳液。

[0791] 由于乳液一般含有可容易地支持微生物生长的大量成分如碳水化合物、蛋白质、固醇及磷脂质,这些制剂一般掺入防腐剂。乳液制剂包括的常用的防腐剂包括对羟基苯甲酸甲酯、对羟基苯甲酸丙酯、季铵盐、氯化苄烷基羟铵、对羟基苯甲酸的酯、及硼酸。抗氧化剂也常常加入乳液制剂中以防止该制剂的变质。所使用的抗氧化剂可以是自由基捕捉剂如生育酚、没食子酸烷基酯、丁基化的羟基茴香醚、丁基化的羟基甲苯、或还原剂如抗坏血酸及偏亚硫酸氢钠、及抗氧化剂增效剂如柠檬酸、酒石酸及卵磷脂。

[0792] 乳液制剂经由护肤途径、口服途径及肠胃外途径的应用及其制造方法已经在文献中回顾(参见,例如,Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems,Allen,LV.,Popovich NG.,and Ansel HC.,2004,Lippincott Williams&Wilkins(8th ed.),New York,NY;Idson,in Pharmaceutical Dosage Forms,Lieberman,Rieger and Banker(Eds.),1988,Marcel Dekker,Inc.,New York,N.Y.,第1卷,第199页)。用于口服递送的乳液制剂因为其容易配制以及在吸收及生物利用度方面的效能而已经广泛使用(参见,例如,Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems,Allen,LV.,Popovich NG.,and Ansel HC.,2004,Lippincott Williams&Wilkins(8th ed.),New York,NY;Rosoff,in Pharmaceutical Dosage Forms,Lieberman,Rieger and

Banker(Eds.),1988,Marcel Dekker,Inc.,New York,N.Y.,第1卷,第245页;Idson,in Pharmaceutical Dosage Forms,Lieberman,Rieger and Banker(Eds.),1988,Marcel Dekker,Inc.,New York,N.Y.,第1卷,第199页)。矿物油基质轻泻剂、油溶性维生素及高脂肪营养制剂属于已经作为o/w乳液而常常口服给药的材料。

[0793] ii.微乳液

[0794] 在本公开的一个实施方案中,RNAi剂及核酸的组合物被配制为微乳液。微乳液可定义为水、油及两性的系统,其是单一光学各向同性且热力学稳定的溶液(参见,例如,Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems,Allen,LV.,Popovich NG.,and Ansel HC.,2004,Lippincott Williams&Wilkins(8th ed.),New York,NY;Rosoff,在Pharmaceutical Dosage Forms,Lieberman,Rieger and Banker(Eds.),1988,Marcel Dekker,Inc.,New York,N.Y.,第1卷,第245页)。典型地,微乳液是通过下述制备的系统,首先,将油分散于表面活性剂水溶液中,随后加入足量的第四成分,通常是中等链长的醇,以形成透明的系统。因此,微乳液也已经描述为两种不互混液体的热力学稳定、各向同性的澄清分散液,该两种液体通过表面活性分子的界面膜予以稳定化(Leung and Shah,在:Controlled Release of Drugs:Polymers and Aggregate Systems,Rosoff,M.,Ed.,1989,VCH Publishers,New York,第185-215页)。微乳液通常经由将三至五种组分组合而制备,这些组分包括油、水、表面活性剂、助表面活性剂及电解质。微乳液是否为油包水(w/o)类型或水包油(o/w)类型取决于所使用的油及表面活性剂的特性,也取决于表面活性剂分子的极性头部及烃类尾部的结构及几何封装(Schott,in Remington's Pharmaceutical Sciences,Mack Publishing Co.,Easton,Pa.,1985,p.271)。

[0795] 使用相图的现象学途径已经广泛研究,且已经使本领域技术人员获得如何配制微乳液的全面知识(参见,例如,Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems,Allen,LV.,Popovich NG.,and Ansel HC.,2004,Lippincott Williams&Wilkins(8th ed.),New York,NY;Rosoff,在Pharmaceutical Dosage Forms,Lieberman,Rieger and Banker(Eds.),1988,Marcel Dekker,Inc.,New York,N.Y.,第1卷,第245页;Block,在Pharmaceutical Dosage Forms,Lieberman,Rieger and Banker(Eds.),1988,Marcel Dekker,Inc.,New York,N.Y.,第1卷,第335页)。与常规乳液相比,微乳液具备将制剂中水不溶性药物溶解为自发形成的热力学稳定的液滴的优点。

[0796] 在微乳剂的制备中使用的表面活性剂包括但不限于,离子性表面活性剂、非离子性表面活性剂、Brij 96、聚氧乙烯油基醚、聚甘油脂肪酸酯、四甘油单月桂酸酯(ML310)、四甘油单油酸酯(MO310)、六甘油单油酸酯(PO310)、六甘油五油酸酯(PO500)、十甘油单癸酸酯(MCA750)、十甘油单油酸酯(MO750)、十甘油倍半油酸酯(SO750)、十甘油十油酸酯(DAO750),单独使用或与助表面活性剂合用。助表面活性剂,一般是短链醇如乙醇、1-丙醇及1-丁醇,用来通过因为在表面活性剂分子间生成的空洞空间而渗透入表面活性剂膜并随后创建失序膜,从而增加界面流动性。然而,微乳剂可不使用助表面活性剂而制备,且不含醇的自乳化微乳液系统是该领域中已知者。水性相典型可以是但不限于,水、药物的水溶液、甘油、PEG300、PEG400、聚甘油、丙二醇、及乙二醇的衍生物。油性相可包括但不限于,材料诸如Captex 300、Captex 355、Capmul MCM、脂肪酸酯、重链(C8-C12)单、二及三甘油酯、

聚氧乙基化的甘油基脂肪酸酯、脂肪醇、聚二醇化的甘油酯、饱和聚二醇化的C8-C10甘油酯、植物油及硅油。

[0797] 由药物溶解性的观点及增强的药物吸收来看，微乳液是尤其感兴趣的。已经提出，基于脂质的微乳液(o/w及w/o两者)增强药物的口服生物利用度，该药物包括肽(参见，例如，美国专利第6,191,105号、第7,063,860号、第7,070,802号、第7,157,099号；Constantinides et al.,Pharmaceutical Research,1994,11,1385-1390;Ritschel,Meth.Find.Exp.Clin.Pharmacol.,1993,13,205)。微乳液提供下列优点：改善的药物溶解性、保护药物不被酶水解、由于引入表面活性剂导致的膜流动性及可透过性的改变造成的可能提升的药物吸收、容易制备、比固体剂型更易口服给药、改善的临床潜能、以及降低的毒性(参见，例如，美国专利第6,191,105号、第7,063,860号、第7,070,802号、第7,157,099号；Constantinides et al.,Pharmaceutical Research,1994,11,1385;Ho et al.,J.Pharm.Sci.,1996,85,138-143)。当在环境温度下将微乳液的组分带至一起时，一般可自发形成微乳液。这在配制热稳定药物、肽或RNAi剂时尤其具有优势。也已经发现，微乳液在化妆品及医药两种应用中活性组分的透皮递送中有效。预期本公开的微乳液组合物及制剂将促进胃肠道对RNAi剂及核酸的系统性吸收，以及改善对RNAi剂及核酸的局部细胞摄取。

[0798] 本公开的微乳液也可含有额外的组分及添加剂如失水山梨醇单硬脂酸酯(Grill 3)、Labrasol、以及渗透增强剂，以改善制剂的特性并增强对本发明的RNAi剂及核酸的吸收。本公开的微乳液中使用的渗透增强剂可归类为属于下述五大类之一：表面活性剂、脂肪酸、胆汁盐、螯合剂、及非螯合非表面活性剂(Lee et al.,Critical Reviews in Therapeutic Drug Carrier Systems,1991,p.92)。这些类型各自已经于上文讨论。

[0799] iii.微粒

[0800] 本公开的RNAi剂可并入颗粒例如微粒中。微粒可通过喷雾干燥生产，但也可通过其他方法生产，这些其他方法包括冻干、蒸发、流动床干燥、真空干燥、或这些技术的组合。

[0801] iv.渗透增强剂

[0802] 在一个实施方案中，本公开采用多种渗透增强剂以实现核酸尤其是RNAi剂至动物皮肤的有效递送。大多数药物以经离子化及未经离子化两种形式存在于溶液中。然而，一般仅脂溶性或亲脂性药物容易地跨越细胞膜。已经发现，如果待被跨越的细胞膜经渗透增强剂处理，则即便是非亲脂性药物仍能够跨越该细胞膜。渗透增强剂除了有助于非亲脂性药物跨越细胞膜的扩散之外，也提升亲水性药物的渗透能力。

[0803] 渗透增强剂可分类为属于下述五大类的之：即，表面活性剂、脂肪酸、胆汁盐、螯合剂、及非螯合非表面活性剂(参见，例如，Malmsten,M.Surfactants and polymers in drug delivery,Informa Health Care,New York,NY,2002;Lee et al.,Critical Reviews in Therapeutic Drug Carrier Systems,1991,p.92)。上述各类别的渗透增强剂更详细描述于下。

[0804] 表面活性剂(或“表面活性试剂”)是化学实体，当其溶解在水性溶液中时，降低该溶液的表面张力或该水性溶液与另一液体间的界面张力，结果为RNAi剂通过黏膜的吸收得以提升。这些渗透增强剂除了胆酸盐及脂肪酸外还包括，例如，月桂基硫酸钠、聚氧乙烯-9-月桂基醚及聚氧乙烯-20-鲸蜡基醚(参见，例如，Malmsten,M.Surfactants and polymers in drug delivery,Informa Health Care,New York,NY,2002;Lee et al.,Critical

Reviews in Therapeutic Drug Carrier Systems,1991,p.92);以及全氟化学乳液如FC-43。Takahashi et al.,J.Pharm.Pharmacol.,1988,40,252)。

[0805] 作为渗透增强剂而起作用的多种脂肪酸及其衍生物包括,例如,油酸、月桂酸、癸酸(正癸酸)、肉豆蔻酸、棕榈酸、硬脂酸、亚油酸、亚麻酸、二癸酸酯、三癸酸酯、甘油单油酸酯(1-单油酰基-rac-甘油)、甘油二月桂酸酯、辛酸、花生油酸、甘油-1-单癸酸酯、1-十二烷基氮杂环庚-2-酮、酰基肉碱、酰基胆碱、其C1-20烷基酯(例如,甲酯、异丙酯及叔丁酯)、及其单甘油酯及二甘油酯(即,油酸酯、月桂酸酯、癸酸酯、肉豆蔻酸酯、棕榈酸酯、硬脂酸酯、亚油酸酯等)(参见,例如,Touitou,E.,et al.Enhancement in Drug Delivery,CRC Press,Danvers,MA,2006;Lee et al.,Critical Reviews in Therapeutic Drug Carrier Systems,1991,p.92;Muranishi,Critical Reviews in Therapeutic Drug Carrier Systems,1990,7,1-33;El Hariri et al.,J.Pharm.Pharmacol.,1992,44,651-654)。

[0806] 胆汁的生理学作用包括促进对脂质及脂溶性微生物的分散及吸收(参见,例如,Malmsten,M.Surfactants and polymers in drug delivery,Informa Health Care,New York,NY,2002;Brunton,第38章,在:Goodman&Gilman's The Pharmacological Basis of Therapeutics,9th Ed.,Hardman et al.Eds.,McGraw-Hill,New York,1996,pp.934-935中)。多种天然胆汁盐及其合成衍生物作为渗透增强剂而起作用。因此,术语"胆汁盐"包括胆汁的任意天然组分以及它们的任意合成衍生物。适宜的胆汁盐包括,例如,胆酸(或其药学可接受的钠盐,胆酸钠)、脱氢胆酸(脱氢胆酸钠)、脱氧胆酸(脱氧胆酸钠)、甘胆酸(甘胆酸钠)、糖胆酸(糖胆酸钠)、糖脱氧胆酸(糖脱氧胆酸钠)、牛磺胆酸(牛磺胆酸钠)、牛磺脱氧胆酸(牛磺脱氧胆酸钠)、鹅脱氧胆酸(鹅脱氧胆酸钠)、乌索脱氧胆酸(UDCA)、牛磺-24,25-二氢-褐霉素钠(STDHF)、糖基二氢褐霉素钠及聚氧乙烯-9-月桂基醚(POE)(参见,例如,Malmsten,M.Surfactants and polymers in drug delivery,Informa Health Care,New York,NY,2002;Lee et al.,Critical Reviews in Therapeutic Drug Carrier Systems,1991,page 92;Swinyard,第39章,在:Remington's Pharmaceutical Sciences,18th Ed.,Gennaro,ed.,Mack Publishing Co.,Easton,Pa.,1990,pages 782-783;Muranishi,Critical Reviews in Therapeutic Drug Carrier Systems,1990,7,1-33;Yamamoto et al.,J.Pharm.Exp.Ther.,1992,263,25;Yamashita et al.,J.Pharm.Sci.,1990,79,579-583)。

[0807] 与本公开关联使用的螯合剂可定义为,通过与金属离子形成复合物而将该金属离子从溶液中移除的化合物,结果为通过黏膜进行的RNAi剂吸收得以提升。关于它们作为渗透增强剂在本公开中的用途,螯合剂具有还作为DNase抑制剂而起作用的附加优点,因为大多数特征化DNA核酸酶需要用于淬火的二价金属离子并因此被螯合剂所抑制(Jarrett,J.Chromatogr.,1993,618,315-339)。适宜的螯合剂包括但不限于,乙二胺四乙酸二钠(EDTA)、柠檬酸、柳酸盐(例如,柳酸钠、5-甲氧基柳酸盐及高香草酸钠)、胶原的N-酰基衍生物、laureth-9及beta-二酮的N-胺基酰基衍生物(烯胺类)(参见,例如,Katdare,A.et al.,Excipient development for pharmaceutical,biotechnology,and drug delivery,CRC Press,Danvers,MA,2006;Lee et al.,Critical Reviews in Therapeutic Drug Carrier Systems,1991,第92页;Muranishi,Critical Reviews in Therapeutic Drug Carrier Systems,1990,7,1-33;Buur et al.,J.Control Rel.,1990,14,43-51)。

[0808] 如本文中所用，非螯合非表面活性剂渗透增强剂化合物可定义为，证明其作为螯合剂或作为表面活性剂的活性不显著但仍然增强通过消化道黏膜进行的RNAi剂吸收的化合物(参见，例如，Muranishi,Critical Reviews in Therapeutic Drug Carrier Systems,1990,7,1-33)。这一类渗透增强剂包括，例如，不饱和环状脲、1-烷基-及1-烯基氮杂环烷酮衍生物(Lee et al.,Critical Reviews in Therapeutic Drug Carrier Systems,1991,page 92)；以及非类固醇抗炎剂如双氯芬酸钠、吲哚美辛(indomethacin)及丁二苯吡唑二酮(Yamashita et al.,J.Pharm.Pharmacol.,1987,39,621-626)。

[0809] 在细胞水平增强RNAi剂摄入的剂也可加入本公开的药物组合物及其他组合物中。例如，阳离子脂质如lipofectin(Junichi等人，美国专利第5,705,188号)、阳离子性甘油衍生物、及聚阳离子性分子诸如聚赖氨酸(WO 97/30731)，也已知提升dsRNA的细胞摄入。

[0810] 其他剂可用以增强所给药的核酸的渗透，包括二醇类如乙二醇及丙二醇、吡咯类如2-吡咯、氮酮类、及萜类如柠檬烯及薄荷酮。

[0811] vi.赋形剂

[0812] 与载剂化合物相比，“药物载剂”或“赋形剂”是药学可接受的溶剂、悬浮剂或其他用于将一种或多种核酸递送至动物的药学惰性囊泡。该赋形剂可以是液体或固体，且当与核酸及给定药物组合物的其他组分合并时，是基于所考虑的计划给药模式而选择，以提供所希望的体积、一致性等。典型的药物载剂包括但不限于，结合剂(例如，预胶凝的玉米淀粉、聚乙烯基吡咯烷酮或羟丙基甲基纤维素等)；填料(例如，乳糖及其他糖类、微晶纤维素、果胶、明胶、硫酸钙、乙基纤维素、聚丙烯酸酯或磷酸氢钙等)；润滑剂(例如，硬脂酸镁、云母、氧化硅、胶体二氧化硅、硬脂酸、金属硬脂酸盐、氢化植物油、玉米淀粉、聚乙二醇、苯甲酸钠、醋酸钠等)；崩解剂(例如，淀粉、淀粉乙醇酸钠等)；以及润湿剂(例如，月桂基硫酸钠等)。

[0813] 不与核酸进行有害反应的适用于非肠胃外给药的药学可接受的有机或无机赋形剂也可用以配制本公开的组合物。适宜的药学可接受的载剂包括但不限于，水、盐溶液、醇类、聚乙二醇、明胶、乳糖、直链淀粉、硬脂酸镁、滑石、硅酸、黏性石蜡、羟甲基纤维素、聚乙烯基吡咯烷酮等。

[0814] 用于核酸的外用给药的制剂可包括无菌及非无菌水性溶液、在常用溶剂如醇类中的非水性溶液、或核酸在液体或固体油基质中的溶液。这些溶液也可含有缓冲剂、稀释剂及其他适宜的添加剂。可使用不与核酸进行有害反应的适用于非肠胃外给药的药学可接受的有机或无机赋形剂。

[0815] 适宜的药学可接受的赋形剂包括但不限于，水、盐溶液、醇、聚乙二醇、明胶、乳糖、直链淀粉、硬脂酸镁、滑石、硅酸、黏性石蜡、羟甲基纤维素、聚乙烯基吡咯烷酮等。

[0816] vii.其他组分

[0817] 本公开的组合物可额外地含有常见于药物组合物中的其他辅助组分，用量为它们在该领域中常用的水平。因此，例如，这些组合物可含有额外、可相容、药学活性的材料诸如，例如，止痒剂、收敛剂、局部麻醉剂或抗炎剂，或可含有可用于物理上配制多种类型的本公开的组合物的材料诸如染料、芳香剂、防腐剂、抗氧化剂、遮光剂、增稠剂及稳定剂。然而，当加入此类材料时，此类材料应不过度干扰本公开的组合物的组分的生物活性。该制剂可经无菌化，且(若需要)与不与该制剂的核酸进行有害反应的佐剂如润滑剂、防腐剂、稳定

剂、润湿剂、乳化剂、用于影响渗透压的盐类、缓冲剂、着色剂、香味剂或芳香物质等混合。

[0818] 水性悬浮液可含有增加悬浮液黏度的物质，该物质包括，例如，羧甲基纤维素钠、山梨醇或葡聚糖。悬浮液也可含有稳定剂。

[0819] 在一些实施方案中，本公开提出的药物组合物包括(a)一种或多种RNAi剂化合物及(b)一种或多种剂，其通过非RNAi机制而发挥功能且有用于治疗C9orf72相关疾患。此类剂的示例包括但不限于，单胺抑制剂、利血平、抗痉挛剂、抗精神病药及抗抑郁药。

[0820] 此类化合物的毒性及治疗功效可通过标准药学过程在细胞培养物或实验动物中测定，如测定LD_{50}(将群体的50％致死的剂量)及ED_{50}(对群体的50％治疗有效的剂量)。毒性与治疗功效间的剂量比率是治疗指数，且其可表示为LD_{50}/ED_{50}的比率。显现高治疗指数的化合物是优选的。

[0821] 从细胞培养测定及动物研究中获得的数据可用于配制在人体内使用的剂量范围。本公开提出的组合物的剂量通常处于包括ED_{50}在内的具低毒性或无毒性的循环浓度范围。剂量可依据所采用的剂型及所使用的给药途径而在此范围内变动。对于在本公开提出的方法中使用的任意化合物，最初可从细胞培养物测定中构建治疗有效的剂量。剂量可在动物模型中配制为实现该化合物或(当适宜时)靶标序列的多肽产物的循环血浆浓度范围(例如，实现多肽的降低的浓度)，该范围包括如在细胞培养物中测定的IC_{50}(即，实现对症状的半最大抑制时该测试化合物的浓度)。此信息可用来更准确地确定可用于人体的剂量。例如，可通过高效液相层析分析术测量血浆中的水平。

[0822] 除了如上文讨论的给药之外，本公开提出的RNAi剂也可与其他已知的在治疗由核苷酸重复序列表达介导的病理进程中有效的剂合用。在任何情况下，主治医生均可基于该领域中已知或本文所述的标准功效测量方法观察的结果而确定RNAi剂给药的量及时机。

[0823] VII.试剂盒

[0824] 在某些方面，本公开提供包括合适容器的试剂盒，该容器含有siRNA化合物例如双链siRNA化合物或ssiRNA化合物(例如，前体，例如，可加工为ssiRNA化合物的较大siRNA化合物，或编码siRNA化合物的DNA，例如，双链siRNA化合物、或ssiRNA化合物或其前体)。

[0825] 此类试剂盒包括一种或多种dsRNA剂及使用说明，例如，用于给药预防或治疗有效量的dsRNA剂的使用说明书。dsRNA剂可处于小瓶内或预填充的注射器内。试剂盒可任选进一步包含用于给药dsRNA剂的工具(例如，注射装置，诸如预填充的注射器)或用于测量对C3的抑制的工具(例如，用于测量对C9orf72 mRNA、C9orf72蛋白质和/或C9orf72活性的抑制的工具)。此类用于测量对C9orf72的抑制的工具可包含用于从受试者获得样本例如SCF和/或血浆样本的工具。本发明的试剂盒可任选进一步包含用于确定治疗有效量或预防有效量的装置。

[0826] 在某些实施方案中，药物制剂的个体组分可提供于一个容器内。另选地，可能希望的是将药物制剂的组分单独提供于两个或更多个容器中，例如，一个用于siRNA化合物制备物的容器以及至少另一个用于载剂化合物的容器。试剂盒可包装成多种不同配置，诸如一个盒子中的一个或多个容器。不同组分可组合，例如，根据试剂盒提供的使用说明书组合。组分可根据本文所述的方法组合，例如以制备并给药一种药物组合物。试剂盒也可包括递送装置。

[0827] VII.抑制C9orf72表达的方法

[0828] 本公开还提供抑制C9orf72基因在细胞内的表达的方法。该方法包括使细胞与其量足以抑制该细胞内C9orf72的表达的RNAi剂例如双链RNAi剂接触，从而抑制该细胞内C9orf72的表达。在本公开的某些实施方案中，C9orf72优先在CNS(例如，脑)细胞中被抑制。

[0829] 细胞与RNAi剂例如双链RNAi剂的接触可在体外或体内进行。使细胞与RNAi剂在体内接触包括使受试者例如人类受试者体内的细胞或细胞群与RNAi剂接触。体外接触细胞的方法与体内接触细胞的方法的组合也是可能的。

[0830] 与细胞接触可以是直接或间接的，如上文讨论。此外，与细胞的接触可经由靶向配体施行，该配体包括本文所描述或该领域中已知的任意配体。在一些实施方案中，靶向配体是碳水化合物部分，例如GalNAc配体，或将RNAi剂引导至感兴趣的位点的任意其他配体。

[0831] 如本文中所用，术语“抑制”与“减轻”、“沉默”、“下调”、“阻抑”及其他类似术语可互换使用，且包括任意水平的抑制。在某些实施方案中，抑制的水平，例如，针对本公开的RNAi剂，可在细胞培养条件下评估，例如，其中，细胞培养物中的细胞经由Lipofectamine™介导的转染以接近10nM或更低、1nM或更低者等的细胞浓度转染。给定RNAi剂的敲低可经由将细胞培养物中处理前的水平与细胞培养物中处理后的水平比较，任选还与经混杂或其他形式的对照RNAi剂平行处理的细胞进行比较而确定。细胞培养物中例如约50％的敲低，可因此被鉴定为“抑制”、“减轻”、“下调”或“阻抑”已经出现的标志。明确预期，在如该领域所描述的适宜控制的条件下，也可于体内系统中评估本公开的RNAi剂所靶向的mRNA或所编码的蛋白质水平(并因此评估由本公开的RNAi剂造成的“抑制”等的程度)。

[0832] 如本文中所用，短语“抑制C9orf72基因的表达”或“抑制C9orf72的表达”包括抑制任意C9orf72基因(例如，小鼠C9orf72基因、大鼠C9orf72基因、猴C9orf72基因、或人类C9orf72基因)以及编码C9orf72蛋白质的C9orf72基因的变体或突变体(例如，在该基因的内含子中具有经扩增的六核苷酸重复序列的C9orf72)的表达。因此，在经基因操纵的细胞、细胞群或生物体的语境中，C9orf72基因可以是野生型C9orf72基因、突变C9orf72基因或转基因C9orf72基因。

[0833] “抑制C9orf72基因的表达”包括对C9orf72的任意水平的抑制，如对C9orf72基因的表达的至少部分的阻抑，例如至少20％的抑制。在某些实施方案中，抑制是至少30％，至少40％，或优选地，至少50％。在其他实施方案中，抑制是不超过50％，例如，不超过50％、45％、40％、35％、30％、25％、20％、15％、10％或5％。

[0834] C9orf72基因的表达可基于与C9orf72基因表达相关的任意变量而评估，例如，C9orf72 mRNA水平(例如，正义mRNA、反义mRNA、总C9orf72 mRNA、含正义C9orf72重复序列的mRNA和/或含反义C9orf72重复序列的mRNA)或C9orf72蛋白质水平(例如，总C9orf72蛋白质、野生型C9orf72蛋白质或含经扩增重复的蛋白质)，或例如含正义或反义的团簇的水平和/或异常二肽重复蛋白质的水平。

[0835] 抑制可通过这些变量的一者或多者的绝对水平或与对照水平相比的相对水平的下降而评估。对照水平可以是该领域中使用的任意类型的对照水平，如给药前的基线水平，或自未治疗或经对照物(例如，仅含缓冲剂的对照物或非活性剂的对照物)治疗的类似受试者、细胞或样本测得的水平。

[0836] 例如，在本公开的方法的一些实施方案中，C9orf72基因的表达(例如，通过含正义或反义的团簇和/或异常二肽重复蛋白质水平评估的)被抑制至少20％、30％、40％、50％、

60％、70％、80％、85％、90％或95％，或被抑制到低于该测定的检测限。在本公开的方法的其他实施方案中，C9orf72基因的表达(例如，通过mRNA或蛋白质表达水平评估的)被抑制不超过50％、45％、40％、35％、30％、25％、20％、15％、10％或5％。在某些实施方案中，该方法包括对C9orf72表达的临床相关抑制，例如，通过在使用剂治疗受试者以降低C9orf72表达之后的临床相关结局证明的。

[0837] C9orf72基因表达的抑制可通过第一细胞或细胞群(此类细胞可存在于来源于受试者的样本中)所表达的mRNA量相对于基本上与第一细胞或细胞群相同但未经这样的处理的第二细胞或细胞群(未经RNAi剂处理或未经靶向感兴趣基因的RNAi剂处理的对照细胞)降低而体现，在该细胞或细胞群中，C9orf72基因被转录并且已经处理(例如，通过使该一个或多个细胞与本公开的RNAi剂接触，或通过将本公开的RNAi剂给药至其体内存在或曾经存在该细胞的受试者)，使得C9orf72基因被抑制。抑制的程度可通过以下表达：

[0838] $$\frac{(\text{对照细胞中的mRNA}) - (\text{经处理细胞中的mRNA})}{(\text{对照细胞中的mRNA})} \bullet 100\%$$

[0839] 在其他实施方案中，对C9orf72基因表达的抑制可在参数降低方面进行评估，该参数在功能上与C9orf72基因表达例如C9orf72蛋白质表达、含正义或反义的团簇和/或异常二肽重复蛋白质的水平相关联。C9orf72基因沉默可在与表达构建体物同源或异源的表达C9orf72的任意细胞内测定，并且通过本领域中已知的任意测定测量。

[0840] C9orf72蛋白表达的抑制可通过细胞或细胞群所表达的C9orf72蛋白水平(或例如本文所述的功能性参数)，例如，于源自受试者的样本中表达的蛋白质水平的降低而体现。如上所述，对于mRNA阻抑的评估，经处理的细胞或细胞群中蛋白质表达水平的抑制可类似地表达为相对于对照细胞或细胞群中蛋白质水平的百分比。

[0841] 可用来评估C9orf72基因表达的抑制的对照细胞或细胞群包括尚未与本公开的RNAi剂接触的细胞或细胞群。例如，对照细胞或细胞群可源自使用RNAi剂治疗受试者之前的个体受试者(例如，人类或动物受试者)。

[0842] 细胞或细胞群体所表达的C9orf72 mRNA水平可使用该领域中已知用于评估mRNA表达的任意方法测定。在一个实施方案中，样本中的C9orf72表达水平可通过检测经转录的聚核苷酸或其蛋白质例如C9orf72基因的mRNA而测定。可使用RNA提取技术，包括例如使用酸酚/异硫氰酸胍提取(RNAzol B;Biogenesis)、RNeasyTM RNA制备试剂盒(Qiagen®)或PAXgene(PreAnalytix,Switzerland)，从细胞中提取RNA。采用核糖核酸杂交的典型测定型式包括核连续测定、RT-PCR、RNase保护测定、北方印渍术、原位杂交及微数组分析。链特异性C9orf72mRNA可使用定向RT-PCR和/或液滴数字PCR方法检测，该方法描述于例如Jiang,et al.supra、Lagier-Tourenne,et al.,supra及Jiang,et al.,supra中。循环C9orf72 mRNA可使用WO2012/177906中所述的方法检测，该专利的全部内容通过援引并入本文。

[0843] 在一些实施方案中，使用核酸探针确定C9orf72的表达水平。如本文所用，术语“探针”指代能够选择性地结合至特异性C9orf72核酸或蛋白质或其片段的任意分子。探针可由本领域技术人员合成，或来源于适宜的生物学制剂。探针可特异性地设计为被标记。可用作探针的分子的示例包括但不限于RNA、DNA、蛋白质、抗体及有机分子。

[0844] 分离的mRNA可用于杂交或扩增测定中，该测定包括但不限于，Southern或

Northern分析、聚合酶链式反应(PCR)分析及探针数组。一种用于测定mRNA水平的方法包括使分离的mRNA与可杂交至C9orf72mRNA的核酸分子(探针)接触。在一个实施方案中,例如,通过使分离的mRNA在琼脂糖凝胶上电泳并将mRNA从该凝胶转移至膜诸如硝基纤维素膜,从而将mRNA固定于固体表面并使其与探针接触。在候选的实施方案中,例如,在Affymetrix®基因芯片数组中,将探针固定于固体表面并使mRNA与该探针接触。熟练人士可轻易地调整已知的mRNA检测方法以用于测定C9orf72 mRNA的水平。

[0845] 用于测定样本中C9orf72表达水平的方法包括例如样本中的mRNA的核酸扩增或逆转录(以制备cDNA)的过程,该过程是例如通过RT-PCR(Mullis于1987年于美国专利4,683,202中详述的实验性实施方案)、连接酶链式反应(Barany(1991)Proc.Natl.Acad.Sci.USA 88:189-193)、自我持续的序列复制(Guatelli et al.(1990)Proc.Natl.Acad.Sci.USA87:1874-1878)、转录扩增系统(Kwoh et al.(1989)Proc.Natl.Acad.Sci.USA 86:1173-1177)、Q-β复制酶(Lizardi et al.(1988)Bio/Technology6:1197)、滚环式复制(Lizardi et al.,US Patent No.5,854,033)或任意其他核酸扩增方法进行,之后使用本领域技术人员熟知的技术检测所扩增的分子。如果核酸分子以非常低的数量存在,则这些检测方法尤其有用于检测此类分子。在本公开的具体方面,通过定量萤光RT-PCR(即,TaqMan™系统)、通过Dual-Glo®萤光素酶测定、或通过其他该领域认可的用于测量C9orf72表达或mRNA水平的方法测定C9orf72表达水平。

[0846] 可使用膜印迹(诸如杂交分析中所用的,诸如northern印迹、southern印迹、点等)或微孔、样本管、凝胶、珠或纤维(或包含经结合的核酸的任意固体支持物)监控C9orf72 mRNA的表达水平。参见,美国专利第5,770,722号、第5,874,219号、第5,744,305号、第5,677,195号及第5,445,934号,这些专利通过援引并入本文。C9orf72表达水平的测定也可包含使用处于溶液中的核酸探针。

[0847] 在一些实施方案中,使用分支DNA(bDNA)测定或实时PCR(qPCR)评估mRNA表达的水平。这一PCR方法的用途描述且例示于本文的实施例中。此类方法也可用于检测C9orf72核酸。

[0848] C9orf72蛋白表达的水平可使用该领域中已知用于测量蛋白质水平的任意方法测定。此类方法包括,例如,电泳、毛细管电泳、高效液相层析术(HPLC)、薄层层析术(TLC)、超扩散层析术、流体或凝胶沉淀反应、吸收光谱、比色测定、分光亮度测定、流式细胞术、免疫扩散(单或双)、免疫电泳、western blotting、放射免疫测定(RIA)、酶联免疫吸附测定(ELISA)、免疫萤光测定、电化学发光测定等。此类测定可用于检测C9orf72蛋白质的存在或复制的蛋白质标志物。

[0849] 含正义或反义的团簇的水平及异常二肽重复蛋白质的水平可使用本领域普通技术人员熟知的方法评估,包括,例如,萤光原位杂交(FISH)、免疫组织化学及免疫测定(参见,例如,Jiang,et al.supra)。在一些实施方案中,本公开的方法在治疗C9orf72相关疾病中的功效通过C9orf72mRNA水平的降低而评估(例如,通过评估CSF样本和/或血浆样本的C9orf72水平、通过脑部生物活检或其他方式)。

[0850] 在本公开的方法的一些实施方案中,将RNAi剂给药至受试者,使得RNAi剂被递送至该受试者体内的具体位点。对C9orf72表达的抑制可使用测量源自受试者体内具体位点例如CNS细胞的样本中C9orf72mRNA(正义mRNA、反义mRNA、总C9orf72 mRNA、含正义C9orf72

重复的mRNA和/或含反义C9orf72重复的mRNA)、C9orf72蛋白质(例如,总C9orf72蛋白质、野生型C9orf72蛋白质或含经扩增重复的蛋白质)、含正义的团簇、含反义的团簇、异常二肽重复蛋白质的水平或水平的改变而评估。在某些实施方案中,该方法包括对C9orf72表达的临床相关抑制,例如,通过在使用剂治疗受试者以降低C9orf72表达之后的临床相关结局证明的,例如,尾状核萎缩(caudate atrophy)的稳定化或抑制(例如,通过容积MRI(vMRI)评估的)、来自受试者的CSF样本中神经元丝轻链(Nfl)水平的稳定化或降低、突变C9orf72 mRNA或经裂解的C9orf72蛋白质的降低(例如,全长度突变C9orf72 mRNA或蛋白质以及经裂解的突变C9orf72 mRNA或蛋白质中的一者或两者)、以及统一C9orf72相关疾病评定量表(Unified C9orf72-associated disease Rating Scale(UHDRS))得分。

[0851] 如本文中所用,术语检测或确定分析物的水平被理解为意指执行这些步骤以确定材料例如蛋白质、RNA是否存在。如本文中所用,检测或确定方法包括检测或确定低于所使用方法的检测水平的分析物水平。

[0852] IX.治疗或预防C9orf72相关疾病的方法

[0853] 如本文所描述,降低含C9orf72重复的RNA可通过RNA干扰的天然过程实现,在该过程中,dsRNA引导靶标RNA被RNA诱导沉默复合物直接裂解,之后RNA裂解片段被细胞核酸酶降解。然而,RNA干扰主要是细胞质过程,预期其不会作用于保留在细胞核中的RNA。含内含子的RNA通常是短寿命,或作为mRNA前体,其被迅速间接为成熟mRNA,或作为被剔除的内含子,于细胞核中迅速降解。因此,可合理地预期,含内含子的RNA将不能用于通过RNA干扰进行靶向。

[0854] 然而,如本文所表明,靶向与GGGGCC重复扩增相邻的内含子序列的siRNA促进含内含子的C9orf72 RNA的累积的降低,同时对C9orf72成熟RNA具有极低效应或无效应。靶向内含子的siRNA也降低二肽重复蛋白质的产生。使人惊奇地,已经发现,所披露的靶向本区域的siRNA在降低含内含子的C9orf72 RNA水平方面有效。结果显示,负责二肽重复蛋白质合成的相当一部分含内含子的C9orf72 RNA位于细胞质中。相比之下,靶向C9orf72 mRNA蛋白质编码序列的siRNA产生mRNA的强烈敲低,但对于含内含子的转录本无效,并且未明显降低二肽重复蛋白质合成。靶向内含子的siRNA与靶向mRNA的siRNA之间的结果发散性表明,两类靶向序列存在于未共价连接的独立RNA上。

[0855] 本文所述的方法提供二肽重复蛋白质合成的治疗性降低,该蛋白质是C9orf72重复扩增疾病的主要致病组分,同时隔离C9orf72 mRNA,从而避免C9orf72蛋白质降低的可能的副作用,而该可能的副作用在使用治疗性策略诸如使用靶向细胞核中的初代C9orf72转录本的反义寡核苷酸时将发生。

[0856] 本文所披露的方法中的一些用于抑制C9orf72靶标RNA于细胞内的表达,该靶标RNA包含六核苷酸重复,该重复包含SEQ ID NO:1的多个连续拷贝。C9orf72靶标RNA可以是,例如,具有致病性六核苷酸重复扩增(具有,例如,至少约30个、至少约35个、至少约40个、至少约50个、至少约60个、至少约70个、至少约80个、至少约100个、至少约200个、至少约300个、至少约400、或至少约500个六核苷酸重复的拷贝)的。此类方法可包括将本文所披露的任意dsRNA剂引入细胞内,从而抑制C9orf72靶标RNA于该细胞内的表达。

[0857] 因此,本公开还提供使用本公开的RNAi剂或含有本公开的RNAi剂的组合物来降低或抑制细胞内C9orf72表达的方法。该方法包括使细胞与本公开的dsRNA接触,并将细胞维

持足以获得C9orf72基因的mRNA转录本降解的时间，从而抑制该C9orf72基因在细胞中的表达。

[0858] 此外，本公开还提供使用本公开的RNAi剂或含有本公开的RNAi剂的组合物来降低细胞内含正义或反义的团簇水平和/或抑制其形成的方法。该方法包括使细胞与本公开的dsRNA接触，从而降低该细胞内含C9orf72正义及反义的团簇的水平。

[0859] 本公开还提供使用本公开的RNAi剂或含有本公开的RNAi剂的组合物来降低细胞内异常二肽重复蛋白质水平和/或抑制其形成的方法。该方法包括使细胞与本公开的dsRNA接触，从而降低该细胞内异常二肽重复蛋白质的水平。

[0860] 此类方法可进一步包括评估C9orf72靶标RNA于该细胞内的表达，和/或评估成熟C9orf72 mRNa于该细胞内的表达。例如，该评估可通过逆转录定量聚合酶链式反应进行，以检测C9orf72靶标RNA。然而，可使用任意其他合适的方法。

[0861] 在本公开的方法中，细胞可在体内或体外接触，即，细胞可处于受试者内。

[0862] 适用于使用本公开的方法治疗的细胞可以是表达C9orf72基因的任意细胞或表达具有扩增GGGGCC重复的C9orf72基因的细胞。适用于在本公开的方法中使用的细胞可以是哺乳动物细胞，例如灵长动物细胞(诸如人类细胞或非人灵长动物细胞，例如猴细胞或黑猩猩细胞)、非灵长动物细胞(诸如大鼠细胞或小鼠细胞)。在一个实施方案中，该细胞是人类细胞，例如，人类CNS细胞。在一些实施方案中，该细胞是非人动物单细胞阶段胚胎、非人动物胚胎干细胞、源自胚胎干细胞的运动神经元、脑细胞、皮质细胞、神经元细胞、肌肉细胞、心脏细胞或生殖细胞。

[0863] 在一些实施方案中，该细胞可包含C9orf72基因座，该基因座包含致病性六核苷酸重复扩增。致病性六核苷酸重复扩增是由大量GGGGCC(SEQ ID NO:1)的重复组成的扩增，其位于分隔基因C9orf72中两个假定的第一非编码外显子(外显子1A及1B)的间隔序列中，与下列致病性读出的一者或两者相关联：(1)可将含有正义及反义重复序列的RNA可视化为在神经元及其他细胞中截然不同的团簇；以及，(2)可于细胞中检测通过重复序列相关的非AUG依赖性翻译从该含正义及反义重复的RNA合成的二肽重复蛋白质，即，聚(甘氨酸-丙氨酸)、聚(谷氨酸-脯氨酸)、聚(甘氨酸-精氨酸)、聚(丙氨酸-脯氨酸)、及聚(脯氨酸-精氨酸)。重复的次数可以是高于正常情况下在来自未患有C9orf72 ALS或C9orf72 FTD的人的基因座中所见的重复的数目。或者，致病性六核苷酸重复扩增可以是来自患有C9orf72 ALS或C9orf72 FTD的受试者的C9orf72基因座中的扩增(即，重复次数)。致病性六核苷酸重复扩增具有GGGGCC(SEQ ID NO:1)的多次重复。例如，致病性六核苷酸重复扩增可具有，例如，至少约30个、至少约35个、至少约40个、至少约50个、至少约60个、至少约70个、至少约80个、至少约100个、至少约200个、至少约300个、至少约400、或至少约500个六核苷酸重复的拷贝。

[0864] 该细胞可以是来自患有C9orf72-六核苷酸重复扩增相关疾病(包括，例如，C9orf72 ALS或C9orf72 FTD)的受试者或处于发展出该疾病风险下的受试者的细胞(例如，神经元或运动神经元)。

[0865] 本文所披露的方法中的细胞可以是任意类型的包含C9orf72基因组的细胞。C9orf72基因组可包含如本文他处所述的六核苷酸重复扩增序列或致病性六核苷酸重复扩增序列。六核苷酸重复扩增序列可包含SEQ ID NO:1中详述的六核苷酸序列的超过100次重

复。

[0866] C9orf72六核苷酸重复扩增序列通常是包含SEQ ID NO:1中详述的六核苷酸序列的至少两个实例(即,两次重复)的核苷酸序列。在一些六核苷酸重复扩增序列中,该重复是连续的(彼此相邻而无间隔序列)。重复扩增序列可定位在例如第一非编码内源性外显子与内源性C9orf72基因座的外显子2之间。

[0867] 六核苷酸重复扩增序列可具有任意次数的重复。例如,该重复扩增序列可包含超过约95次重复、超过约96次重复、超过约97次重复、超过约98次重复、超过约99次重复、超过约100次重复、超过约101次重复、超过约102次重复、超过约103次重复、超过约104次重复、超过约105次重复、超过约150次重复、超过约200次重复、超过约250次重复、超过约295次重复、超过约296次重复、超过约297次重复、超过约298次重复、超过约299次重复、超过约300次重复、超过约301次重复、超过约302次重复、超过约303次重复、超过约304次重复、超过约305次重复、超过约350次重复、超过约400次重复、超过约450次重复、超过约500次重复、超过约550次重复、超过约595次重复、超过约596次重复、超过约597次重复、超过约598次重复、超过约599次重复、超过约600次重复、超过约601次重复、超过约602次重复、超过约603次重复、超过约604次重复、或超过约605次重复。或者,该重复扩增序列可包括至少约95次重复、至少约96次重复、至少约97次重复、至少约98次重复、至少约99次重复、至少约100次重复、至少约101次重复、至少约102次重复、至少约103次重复、至少约104次重复、至少约105次重复、至少约150次重复、至少约200次重复、至少约250次重复、至少约295次重复、至少约296次重复、至少约297次重复、至少约298次重复、至少约299次重复、至少约300次重复、至少约301次重复、至少约302次重复、至少约303次重复、至少约304次重复、至少约305次重复、至少约350次重复、至少约400次重复、至少约450次重复、至少约500次重复、至少约550次重复、至少约595次重复、至少约596次重复、至少约597次重复、至少约598次重复、至少约599次重复、至少约600次重复、至少约601次重复、至少约602次重复、至少约603次重复、至少约604次重复、或至少605次重复。在一具体实施例中,六核苷酸重复扩增序列包含超过约100次重复、超过约300次重复、超过约600次重复,至少约100次重复、至少约300次重复或至少约600次重复。

[0868] 在非人细胞的细胞中,重复扩增序列可以是异源重复扩增序列。或者,重复扩增序列可以是内源性重复扩增序列。当于核酸语境中使用时,术语异源指示该核酸包含自然条件下不会一起出现于相同分子中的至少两个链段(segment)。例如,当参照核酸的链段而使用时,术语异源指示该核酸包括两个或更多个子序列,这些子序列于自然条件下未发现彼此间的相同关系(例如,连接在一起)。作为一个示例,核酸的异源区域是位于另一核酸分子内或附接至另一核酸分子的核酸链段,该另一核酸分子于自然条件下未发现与其他分子相关联。例如,核酸的异源区域可包括其侧翼具有自然条件下未发现与C9orf72序列相关联的重复扩增序列的C9orf72序列。或者,核酸的异源区域可包括其侧翼具有内源性非人核酸序列的人核酸序列。

[0869] 在一些实施例中,经修饰的内源性非人类C9orf72基因座也可包含额外的种间同源C9orf72序列。作为一个示例,非人类C9orf72基因座可包含人类C9orf72核苷酸序列。例如,非人动物C9orf72基因座可包含将内源性C9orf72基因座的5'未翻译的和/或非编码内源性非人序列替换为六核苷酸重复扩增序列并且其侧翼具有种间同源人类C9orf72序列

(即,位于如SEQ ID NO:1详述的六核苷酸序列重复的侧翼的异源(例如,人)序列)。异源重复扩增序列和/或侧翼序列可以是天然出现的基因组序列(例如,天然出现的人类基因组序列)。作为具体实施例,跨越内源性外显子1(例如,外显子1A和/或1B)与内源性非人类C9orf72基因座的ATG起始密码子之间(且任选涵盖其至少一部分)的未翻译的和/或非编码序列或其部分,可替换为异源六核苷酸重复扩增序列,且该异源六核苷酸重复扩增序列任选与侧翼种间同源人类C9orf72序列一起。例如,插入内源性C9orf72基因座内的序列可包含,从5'到3':第一异源六核苷酸侧翼序列、SEQ ID NO:1中详述的六核苷酸的多个重复、以及第二异源六核苷酸侧翼序列。

[0870] 在一个实施方案中,将跨越(且任选涵盖)人类C9orf72基因的外显子1A和/或外显子1B的全部或部分的异源人类C9orf72序列插入内源性C9orf72基因座中。例如,第一异源六核苷酸侧翼序列可包含如SEQ ID NO:46详述的序列或其部分,和/或第二异源六核苷酸侧翼序列可包含如SEQ ID NO:47详述的序列或其部分。据此,在一些实施例中,内源性六核苷酸重复扩增序列被插入其中的内源性C9orf72基因座可包含人序列,该人序列包含SEQ ID NO:46和/或SEQ ID NO:47中详述的序列。

[0871] 在另一具体例实施例中,内源性C9orf72基因座包含将小鼠C9orf72基因座的非编码序列的一部分替换为放置为与小鼠C9orf72启动子和/或人调控元件可操作连接的异源人六核苷酸重复扩增序列(例如,可见于人类C9orf72基因的外显子1A和/或1B中的那些)。参见,US2018/0094267及WO 2018/064600,其各自通过援引而出于全部目的整体并入本文。对于插入内源性非人类C9orf72基因座内的目的,异源六核苷酸重复扩增序列可包含多个(即,至少两个实例(重复))如SEQ ID NO:1详述的六核苷酸序列,并且可与跨越(且任选包括)人类染色体9开放阅读框72(C9orf72)的非编码外显子1A及1B的基因组核酸序列或其部分相同或基本上相同。

[0872] 细胞可以是体外、间接体内或体内的。例如,细胞可以是动物体内的体内细胞。细胞或动物可以是雄性或雌性。细胞或动物对于插入内源性C9orf72基因座处的六核苷酸重复扩增序列可以是杂合或纯合。二倍体生物于每个遗传基因座处具有两个等位基因。每对等位基因表示具体遗传基因座的表型。如果在特定基因座处存在两个相同的等位基因,则表型描述为纯合;而如果两个等位基因不同,则描述为杂合。非人动物可包含插入在其种系基因组中内源性C9orf72基因座处的异源六核苷酸重复扩增序列。

[0873] 细胞中的C9orf72表达(例如,通过正义mRNA、反义mRNA、总C9orf72 mRNA、含正义C9orf72重复的mRNA、含反义C9orf72重复的mRNA水平、总C9orf72蛋白质、和/或含C9orf72重复的蛋白质而评估)被抑制约20%、25%、30%、35%、40%、45%或50%。在优选实施方案中,C9orf72表达被抑制不超过50%,例如,不超过50%、45%、40%、35%、30%、25%、20%、15%、10%或5%。

[0874] C9orf72靶标RNA中表达的降低可以是任意量。细胞中的抑制(如通过含正义或反义的团簇和/或异常二肽重复蛋白质水平评估者)是被抑制至少20%、30%、40%,优选至少50%、60%、70%、80%、85%、90%或95%,或被抑制到低于该测定的检测水平。

[0875] 在一些实施方案中,dsRNA剂可将C9orf72靶标RNA诸如包含六核苷酸重复的C9orf72靶标RNA的表达抑制至少约10%、至少约20%、至少约30%、至少约40%、至少约50%、至少约60%、至少约70%、至少约80%、或至少约90%(或抑制到C9orf72靶标RNA无法

检测的点)。例如,这些水平的抑制可于给药至表达包含六核苷酸重复序列的C9orf72靶标RNA的细胞之后约1天内、约2天内、约3天内、约4天内、约5天内、约6天内、约一周内或约24小时至约48小时内实现。降低可以是,例如,相对于以dsRNA剂处理之前的细胞或相对于未经dsRNA剂处理的对照细胞。

[0876] 在一些方法中,本发明的dsRNA剂相对于成熟C9orf72信使RNA的表达而选择性地抑制包含C9orf72靶标RNA诸如六核苷酸重复的C9orf72靶标RNA的表达。本语境中,成熟C9orf72信使RNA是已经间接及加工的C9orf72 RNA转录本。成熟C9orf72信使RNA唯一地由外显子组成并且内含子全部已经移除。在将dsRNA剂给药至表达C9orf72靶标RNA的细胞后,如果C9orf72靶标RNA的表达的降低大于成熟C9orf72信使RNA的表达的相对降低,则dsRNA剂相对于成熟C9orf72信使RNA的表达而选择性地抑制包含内含子六核苷酸重复的C9orf72靶标RNA的表达。例如,在某些实施方案中,本发明的dsRNA剂将成熟C9orf72信使RNA的表达抑制少于约50%、少于约40%、授予约30%、少于约20%、少于约10%或少于约5%(或者,例如,对于表达不具有任何统计学显著或功能上显著的效应)。例如,这些水平的抑制可于给药至表达包含六核苷酸重复序列的C9orf72靶标RNA的细胞之后约1天内、约2天内、约3天内、约4天内、约5天内、约6天内、约一周内或约24小时至约48小时内实现。降低可以是,例如,相对于以dsRNA剂处理之前的细胞或相对于未经dsRNA剂处理的对照细胞。

[0877] 本文所披露的一些方法用于降低细胞内的二肽重复蛋白质合成或二肽重复蛋白质集聚的方法。此类方法可包括将本文所披露的任意的dsRNA剂引入细胞中,从而降低该细胞内的二肽重复蛋白质合成或二肽重复蛋白质集聚。

[0878] 此类方法可进一步包括评估二肽重复蛋白质集聚体(例如,聚(甘氨酸-丙氨酸)、聚(甘氨酸-脯氨酸)、聚(甘氨酸-精氨酸)、聚(丙氨酸-脯氨酸)及聚(脯氨酸-精氨酸))在该细胞中的存在。在具体实施例中,该二肽重复蛋白质可以是聚(甘氨酸-丙氨酸)和/或聚(甘氨酸-脯氨酸)。该评估可例如通过免疫组织化学或western blot分析进行,以检测二肽重复蛋白质集聚体。然而,可使用任意其他合适的方法。

[0879] 二肽重复蛋白质合成或二肽重复蛋白质集聚体的降低可以是任意量。例如,dsRNA剂可将二肽重复蛋白质合成或二肽重复蛋白质集聚体降低至少约10%、至少约20%、至少约30%、至少约40%、至少约50%、至少约60%、至少约70%、至少约80%、或至少约90%(或降低到二肽重复蛋白质集聚体不可检测的点)。例如,这些水平的抑制可于给药至表达包含六核苷酸重复序列的C9orf72靶标RNA的细胞之后约1天内、约2天内、约3天内、约4天内、约5天内、约6天内、约一周内或约24小时至约48小时内实现。降低可以是,例如,相对于以dsRNA剂处理之前的细胞或相对于未经dsRNA剂处理的对照细胞。

[0880] 此类方法可进一步包括评估细胞内细胞核和/或细胞质正义和/或反义C9orf72 RNA团簇的存在。

[0881] 细胞核和/或细胞质正义和/或反义C9orf72 RNA团簇的存在的降低可以是任意量。例如,dsRNA剂可将细胞核和/或细胞质正义和/或反义C9orf72 RNA团簇的存在降低至少约10%、至少约20%、至少约30%、至少约40%、至少约50%、至少约60%、至少约70%、至少约80%、或至少约90%(或降低到细胞核和/或细胞质正义和/或反义C9orf72 RNA团簇无法检测的点)。例如,这些水平的抑制可于给药至表达包含六核苷酸重复序列的C9orf72靶标RNA的细胞之后约1天内、约2天内、约3天内、约4天内、约5天内、约6天内、约一周内或约24

小时至约48小时内实现。降低可以是，例如，相对于以dsRNA剂处理之前的细胞或相对于未经dsRNA剂处理的对照细胞。

[0882] 本公开的体内方法可包括对受试者给药含有RNAi剂的组合物，其中该RNAi剂包括与待治疗的哺乳动物的C9orf72基因的RNA转录本的至少一部分互补的核苷酸序列。当待治疗的有机体是哺乳动物例如人时，该组合物可通过该领域中已知的任意手段给药，该手段包括但不限于，口服、腹膜内或肠胃外途径，包括颅内(例如，脑室内、脑实质内及鞘内)、静脉内、肌肉内、皮下、透皮、气管(气溶胶)、鼻内、直肠内及外用(包括口含及舌下)给药。在某些实施方案中，该组合物是通过静脉内输液或注射而给药。在某些实施方案中，该组合物是通过皮下注射而给药。在某些实施方案中，该组合物是通过鞘内注射而给药。

[0883] 在一些实施方案中，该给药是经由积存注射(depot injection)进行。积存注射可在延长的时间段内以一致的途径释放RNAi剂。因此，积存注射可降低获得所欲效果如所欲的C9orf72抑制或治疗性或预防性效果所需的给药频次。积存注射亦可提供更为一致的血清浓度。积存注射可包括皮下注射或肌肉注射。在优选的实施方案中，积存注射是皮下注射。

[0884] 在一些实施方案中，该给药是经由泵进行。该泵可以是外部泵或经外科手术植入的泵。在另一实施方案中，该泵是经皮下植入的渗透泵。在其他实施方案中，该泵是输液泵。输液泵可用于颅内、静脉内、皮下、动脉内或硬膜内输液。在优选的实施方案中，该输液泵是皮下输液泵。在其他实施方案中，该泵是经外科手术植入的将RNAi剂递送至CNS的泵。

[0885] 可基于局部治疗或系统性治疗是否是所希望的并基于待治疗的面积而选择给药模式。可选择给药的途径及位点以增强靶向性。

[0886] 一方面，本公开还提供抑制C9orf72基因在哺乳动物体内的表达的方法。这些方法是包括对该哺乳动物给药包含靶向该哺乳动物体内细胞中的C9orf72基因的dsRNA的组合物，从而抑制该细胞内的C9orf72基因的表达。基因表达的降低可通过该领域中已知的任意方法及通过本文中所述的方法例如qRT-PCR评估。蛋白质生产的降低可通过该领域中已知的任意方法及通过本文中所述的方法例如ELISA评估。在一个实施方案中，CNS生物活检样本或脑脊液(CSF)样本作为组织材料用于监控C9orf72基因或蛋白质水平(或其替代者)的降低。

[0887] 本公开进一步提供治疗有此需要的受试者的方法。本公开的治疗方法包括将本公开的RNAi剂给药至受试者例如将会受益于抑制C9orf72表达的受试者诸如在C9orf72基因的内含子中具有GGGGCC经扩增核苷酸重复的受试者，给药量为靶向C9orf72基因的RNAi剂或包含靶向C9orf72基因的RNAi剂的药物组合物的治疗有效量。

[0888] 此外，本公开提供预防、治疗或抑制受试者的C9orf72相关疾病或疾患(例如，C9orf72相关的疾患)进展的方法。该方法包括向受试者给药治疗有效量的本文提供的任意RNAi剂例如dsRNA剂或药物组合物，从而预防、治疗或抑制该受试者的C9orf72相关疾病或疾患的进展。

[0889] 在一些实施方案中，该方法用于治疗罹患C9orf72-六核苷酸重复扩增相关疾病、病症或疾患的受试者。此类方法也可用于预防或缓解患有将会受益于包含SEQ ID NO:1的多个连续拷贝的六核苷酸重复的C9orf72靶标RNA表达降低的疾病、疾患或病症的受试者(例如，患有C9orf72-六核苷酸重复扩增相关疾病、病症或疾患或处于发展出该疾病、病症

或疾患风险下的受试者)的至少一种症状的方法。C9orf72靶标RNA可以是,例如,具有致病性六核苷酸重复扩增(具有,例如,至少约30个、至少约35个、至少约40个、至少约50个、至少约60个、至少约70个、至少约80个、至少约100个、至少约200个、至少约300个、至少约400、或至少约500个六核苷酸重复的拷贝)。C9orf72-六核苷酸重复扩增相关疾病、病症或疾患是由C9orf72基因的5'非编码部分中六核苷酸重复序列(GGGGCC;SEQ ID NO:1)的扩增造成或与其相关联的。示例包括肌萎缩侧索硬化(ALS)及额颞叶痴呆(FTD)。与FTD和/或ALS相关的体征或症状包括但不限于,RNA团簇的重复序列长度依赖性形成、特异性RNA结合蛋白质的隔离、以及神经元中重复相关非AUG(AUG)翻译所致的二肽重复蛋白质(例如,聚(甘氨酸-丙氨酸)、聚(甘氨酸-脯氨酸)、聚(甘氨酸-精氨酸)、聚(丙氨酸-脯氨酸)及聚(脯氨酸-精氨酸))的累积及集聚。本发明的dsRNA剂可用于对FTD和/或ALS相关的体征或症状进行治疗性治疗和/或预防的方法中,体征或症状包括但不限于,运动神经元疾病的体征或症状以及痴呆的体征或症状。运动神经元疾病的体征及症状包括,例如,绊倒、掉落物品、胳膊和/或腿的不正常疲劳、口齿不清、肌肉痉挛及抽搐、笑或哭不受控制以及呼吸困难。痴呆的体征及症状可包括,例如,行为改变、性格改变、言语及语言问题及行动相关问题。

[0890] 本公开的RNAi剂可作为"游离RNAi剂"给药。游离RNAi剂是于药物组合物的不存在下给药。裸RNAi剂可处于合适的缓冲溶液中。缓冲溶液可包含醋酸盐、柠檬酸盐、醇溶谷蛋白、碳酸盐、或磷酸盐、或其任意组合。在一个实施方案中,缓冲溶液是磷酸盐缓冲盐水(PBS)。含有RNAi剂的缓冲溶液的pH及渗透压可经调节,使得其适用于给药至受试者。

[0891] 或者,本公开的RNAi剂可作为药物组合物给药,例如作为dsRNA脂质体制剂给药。

[0892] 将会受益于C9orf72基因表达的降低或抑制的受试者是患有C9orf72相关疾病例如C9orf72相关疾病的那些。示例性C9orf72相关疾病包括但不限于,ALS、FTD、C9ALS/FTD以及由于C9orf72扩增引起的亨廷顿样综合征,例如,在C9orf72基因的内含子中具有经扩增的GGGGCC六核苷酸重复的受试者。

[0893] 本公开进一步提供使用RNAi剂及其药物组合物的方法,例如,用于治疗将会受益于C9orf72表达的降低或抑制的受试者,如患有C9orf72相关疾患的受试者,该方法与其他药品或其他治疗方法合用,例如,与已知的药品或已知的治疗方法如那些当下用于治疗这些疾患者合用。例如,在某些实施方案中,靶向C9orf72的RNAi剂与例如如本文中他处所描述或如该领域中以其他方式已知的可用于治疗C9orf72相关病变的剂合用。例如,适用于治疗将会受益于C9orf72表达降低的受试者例如患有C9orf72相关病变的受试者的附加剂可包括当下用于治疗C9orf72相关疾病症状的剂。RNAi剂及附加治疗剂可同时给药和/或在同一组合中给药,例如鞘内给药,或附加治疗剂可作为独立组合物的一部分或在独立的时间或通过该领域中已知或本文中所述的另一方法给药。

[0894] 示例性的附加治疗剂包括,例如,单胺抑制剂例如四苯喹嗪(tetrabenazine(Xenazine))、氘代四苯喹嗪(deutetrabenazine(Austedo))及利血平;抗痉挛剂例如丙戊酸(丙戊酸钠(Depakote)、帝拔癫、Depacon)及可那氮平(Klonopin);抗精神病药例如利培酮(Risperdal)及氟派醇(Haldol);以及抗抑郁药例如帕罗西汀(Paxil)。

[0895] 在一个实施方案中,该方法包括给药本文中提出的组合物,使得靶标C9orf72基因的表达降低,该降低持续至少一个月。在优选的实施方案中,表达降低持续至少2个月、3个月或6个月。

[0896] 优选地，可用于本文中提出的方法及组合物的RNAi剂特异性地靶向靶标C9orf72基因的RNA(原代或经加工的)。使用RNAi剂抑制这些基因的表达的组合物及方法可如本文中所描述的制备及实施。

[0897] 根据本公开的方法给药dsRNA可导致患有C9orf72相关病变的患者体内这些疾病或病变的严重性、体征、症状或标志物的降低。这一语境中，“降低”意指此水平的统计学显著或临床显著的下降。降低可以是，例如，至少5％、10％、15％、20％、25％、30％、35％、40％、45％、50％、55％、60％、65％、70％、75％、80％、85％、90％、95％、或约100％。

[0898] 疾病的治疗或预防的效力可通过下述评估，例如，测量疾病进展、疾病缓解、症状严重性、疼痛的降低、生命质量、持续治疗效果所需的药物剂量、疾病标记物或任何适用于给定的待治疗疾病或预防的靶向的其他可测量的参数。通过测量此类参数的任一者或参数的任意组合而监控治疗或预防的效率，完全处于本领域技术人员的能力范围内。例如，C9orf72相关病变的治疗功效可例如通过对受试者的周期性监控而评估。将后来的读数与最初的读数比较，对医师提供治疗是否有效的指示。通过测量此类参数的任一者或参数的任意组合而监控治疗或预防的效率，完全处于本领域技术人员的能力范围内。与靶向C9orf72的RNAi剂或其药物组合物的给药相关联，“有效对抗”C9orf72相关病变指示，以临床上适宜的模式给药在至少统计学显著分数的患者中导致有益效果，例如症状的改善、治愈、疾病的降低、生命延长、生命质量的改善、或其他通常被熟悉治疗C9orf72相关疾患及相关病因的医生认为积极的效果。

[0899] 当疾病状态的一个或多个参数存在统计学显著的改善时，或预期会恶化或发展出症状者没有恶化或发展出症状时，证明治疗性或预防性效果。作为示例，可测量的疾病参数的至少10％且优选至少20％、30％、40％、50％或更高的有利改变，可以是有效治疗的指示。给定RNAi剂药物或该药物的制剂的效力也可使用该领域中已知的用于给定疾病的实验动物模型来判断。当使用试验动物模型时，当观察到标志物或症状的统计学显著的降低时，则证明治疗的功效。

[0900] 另选地，可通过诊断领域的技术人员基于临床上接受的疾病严重程度分级量表确定的疾病严重程度的降低，测量该功效。任意导致例如使用适宜量表测量的疾病严重程度变小的正向改变，代表使用本文所述的RNAi剂或RNAi制剂进行了足够的治疗。

[0901] 受试者可给药治疗量的dsRNA，诸如约0.01mg/kg至约200mg/kg。

[0902] RNAi剂可经鞘内给药，经静脉内注射给药，或通过在一段时间内定期静脉输注而给药。在某些实施方案中，在初始的治疗方案后，该治疗的实施频次可降低。RNAi剂的给药可降低C9orf72水平，例如，在该患者的细胞、组织、血液、CSF样本或其他区室内。在优选的实施方案中，RNAi剂的给药可降低C9orf72水平，例如，在该患者的细胞、组织、血液、CSF样本或其他区室内，C9orf72水平降低不超过50％。

[0903] 在给药全剂量的RNAi剂之前，可对患者给药较小的剂量诸如5％输液反应，并监控副作用诸如过敏反应。在另一示例中，可监控患者的不希望的免疫刺激效应如增加的细胞因子(例如，TNF-α或INF-α)水平。

[0904] 另选地，RNAi剂可经皮下给药，即，通过皮下注射给药。一次或多次注射可用来将所希望的例如月剂量的RNAi剂递送至受试者。该注射可在一段时间内重复实施。该给药可定期重复实施。在某些实施方案中，于初始的治疗方案后，该治疗的实施频次可降低。重复

剂量方案可包括规则地给药治疗量的RNAi剂，诸如每个月一次或延长至每季一次、每年两次、每年一次。在某些实施方案中，RNAi剂是大约每个月给药一次至大约每季给药一次(即，大约每三个月给药一次)。

[0905] 除非另外定义，否则本文中使用的所有科技术语具有与本发明所属领域普通技术人员所一般理解的相同的含义。尽管在本发明提出的RNAi剂及方法的实践或测试中可使用与本文中所述的方法及材料类似或等同物，但适宜的方法及材料描述如下。本文中提及的所有出版物、专利申请及其他参考文献通过援引而以其整体并入本文。若有矛盾之处，则以包括定义在内的本说明书为准。此外，材料、方法及实施例仅做例示性说明的用而非意图进行限制。

[0906] 序列表与本文一起递交形成所递交的说明书的一部分。

[0907] 所附序列表中列述的核苷酸及氨基酸序列是使用核苷酸碱基的标准字母缩写、氨基酸的三字母代码显示。核苷酸序列遵循在序列的5’端开始并向前行进(即，每一行从左至右)到3’端的标准惯例。如果显示每个核苷酸序列的一条链，则互补链理解为包括在所显示链的任何参考中。当提供编码氨基酸序列的核苷酸序列时，应理解还提供其编码相同氨基酸序列的密码子简并变体。氨基酸序列遵循在序列的氨基端开始并向前行进(即，每一行从左至右)到羧基端的标准惯例。下表3提供对某些本文所披露的序列的简要说明。

[0908] 表3.某些序列的描述

[0909]

SEQ ID NO	类型	描述
1	DNA	GGGGCC C9orf72 六核苷酸重复
2-25	RNA	C9orf72 siRNA 正义及反义链
26	DNA	智人（*Homo sapiens*） C9orf72 mRNA (NM_018325.4)
27	蛋白质	智人 C9orf72 氨基酸 (NP_060795.1)
28	DNA	智人 C9orf72 mRNA (NM_001256054.2)
29	蛋白质	智人 C9orf72 氨基酸(NP_001242983.1)
30	DNA	5’ gRNA-1 靶标序列
31	DNA	5’ gRNA-2 靶标序列
32	DNA	5’ gRNA-3 靶标序列
33	DNA	5’ gRNA-4 靶标序列
34	DNA	3’ gRNA-1 靶标序列
35	DNA	3’ gRNA-2 靶标序列
36	DNA	3’ gRNA-3 靶标序列
37	DNA	3’ gRNA-4 靶标序列

[0910]

SEQ ID NO	类型	描述
38	DNA	锁核酸 (LNA) 正义 RNA 探针
39	DNA	锁核酸 (LNA) 反义 RNA 探针
40	DNA	小鼠（*Mus musculus*） C9orf72 mRNA (NM_001081343.1)
41	蛋白质	小鼠 C9orf72 氨基酸 (NP_001074812.1)
42	DNA	大鼠（*Rattus norvegicus*） C9orf72 mRNA (NM_001007702.1)
43	蛋白质	大鼠 C9orf72 氨基酸(NP_001007703.1)
44	DNA	DNA 正义探针
45	DNA	DNA 反义探针
46	DNA	人 5’ 侧翼序列
47	DNA	人 3’ 侧翼序列
48	DNA	来自图 2 的序列 1
49	DNA	来自图 2 的序列 2
50	DNA	来自图 2 的序列 3
51	DNA	来自图 2 的序列 4
52	DNA	来自图 3 的 gRNA_5 侧_重复 2
53	DNA	来自图 3 的 gRNA_3 侧_重复 1
54	DNA	来自图 3 的 gRNA_3 侧_重复 3
55	DNA	来自图 3 的 gRNA_3 侧_重复 4
56	RNA	5’ gRNA-1 crRNA
57	RNA	5’ gRNA-2 crRNA
58	DNA	来自图 3 的 5’靶基因座
59	DNA	来自图 3 的 3’靶基因座
60	DNA	来自图 6A 及 6B 的序列
61	DNA	分析 A – 正向引物
62	DNA	分析 A – 反向引物
63	DNA	分析 A – 探针
64	DNA	分析 B – 正向引物
65	DNA	分析 B –反向引物
66	DNA	分析 B – 探针
67	DNA	分析 D – 正向引物
68	DNA	分析 D –反向引物
69	DNA	分析 D – 探针
70	DNA	分析 G – 正向引物
71	DNA	分析 G –反向引物
72	DNA	分析 G – 探针
73	DNA	分析 H – 正向引物
74	DNA	分析 H –反向引物
75	DNA	分析 H – 探针
76	RNA	5’ gRNA-3 crRNA
77	RNA	5’ gRNA-4 crRNA
78	RNA	3’ gRNA-1 crRNA
79	RNA	3’ gRNA-2 crRNA
80	RNA	3’ gRNA-3 crRNA

[0911]

SEQ ID NO	类型	描述
81	RNA	3' gRNA-4 crRNA
82	RNA	TracrRNA v2
83	RNA	GGGGU G-富集 RNA 重复
84	RNA	5'-gRNA-1 导引序列
85	RNA	5'-gRNA-2 导引序列
86	RNA	5'-gRNA-3 导引序列
87	RNA	5'-gRNA-4 导引序列
88	RNA	3'-gRNA-1 导引序列
89	RNA	3'-gRNA-2 导引序列
90	RNA	3'-gRNA-3 导引序列
91	RNA	3'-gRNA-4 导引序列
92	DNA	2-正向引物
93	DNA	2-反向引物
94	DNA	智人 C9orf72 mRNA (NM_145005.6)
95	DNA	分析 C – 正向引物
96	DNA	分析 C – 反向引物
97	DNA	分析 C – 探针
98	DNA	分析 E – 正向引物
99	DNA	分析 E – 反向引物
100	DNA	分析 E – 探针
101	DNA	分析 F – 正向引物
102	DNA	分析 F – 反向引物
103	DNA	分析 F – 探针
104	DNA	分析 I – 正向引物
105	DNA	分析 I – 反向引物
106	DNA	分析 I – 探针
107	DNA	分析 J – 正向引物
108	DNA	分析 J – 反向引物
109	DNA	分析 J – 探针
110	蛋白质	智人 C9orf72 氨基酸 (NP_659442.2)
111	RNA	外显子 1A 起始位点至外显子 2 起始位点
112	RNA	反向互补-外显子 1A 起始位点至外显子 2 起始位点(RNA)
113	RNA	包含 3 个完整 GGGCC 重复的外显子 1A 至外显子 1B 之间的内含子序列(RNA)
114	RNA	反向互补-包含 3 个完整 GGGCC 重复的外显子 1A 至外显子 1B 之间的内含子序列(RNA)
115	RNA	外显子 1A 及第一六核苷酸重复之间的内含子序列(RNA)
116	RNA	反向补体-外显子 1A 及第一个六核苷酸重复之间的内含子序列(RNA)
117	RNA	外显子 1A 及重复之间的片段 (RNA)
118	RNA	反向互补- 外显子 1A 及重复之间的片段(RNA)

[0912] 实施例

[0913] 实施例1:C9orf72基因座处的六核苷酸重复扩增

[0914] I.概述

[0915] 肌萎缩侧索硬化(ALS)及额颞叶痴呆(FTD)是毁灭性神经退行性疾病,其在ALS情况下造成运动神经元疾病并且在FTD情况下造成痴呆。两者都是不可改变地致命性的。ALS及FTD可作为自发性疾病或家族性(即,遗传性)疾病存在。ALS及FTD的最常见遗传病因是C9orf72基因的5'非编码部分中六核苷酸重复(GGGGCC;SEQ ID NO:1)的扩增,该基因编码其功能尚未完全得以理解的蛋白质。未受影响的人通常已经在其C9orf72基因中具有数个

及数十个六核苷酸重复,而那些发展出ALS及FTD者从其父母之一遗传了数百至数千个六核苷酸重复拷贝的重复扩增。遗传性观察结果表明,C9orf72 ALS及FTD是显性遗传疾病并且导致病理学功能的增益。

[0916] C9orf72六核苷酸重复扩增如何造成运动神经元疾病及痴呆尚未可知,但C9orf72 ALS及FTLD患者的两种普遍尸检病理学发现与重复扩增相关联:(1)通过萦光原位杂交,可将含有正义及反义重复的RNA可视化为在神经元及其他细胞中截然不同的团簇;以及,(2)通过重复相关的非AUG依赖性翻译从该含正义及反义重复的RNA合成的二肽重复蛋白质,即,聚(甘氨酸-丙氨酸)、聚(谷氨酸-脯氨酸)、聚(甘氨酸-精氨酸)、聚(丙氨酸-脯氨酸)、及聚(脯氨酸-精氨酸),可通过免疫组织化学于细胞中检测。一种疾病假说提出,可视化为团簇的含重复的RNA通过隔离RNA结合蛋白质而破坏细胞RNA代谢。第二种疾病假说认为,二肽重复蛋白对RNA代谢、蛋白质稳态及核细胞质转运具有广泛的毒性效应。两种病理学机制都可能对疾病有所贡献。如果含有C9orf72重复的RNA转录本或其自身或作为二肽重复蛋白质翻译模板而促进ALS及FTLD中的致病性,则常规治疗策略将是摧毁含GGGGCC重复的RNA或根除其被翻译为二肽重复蛋白质的能力。

[0917] C9orf72基因从两个转录起始位点产生转录本。上游位点以候选的非编码外显子1A启动转录,而下游位点以候选的外显子1B启动转录。外显子1A及1B两者都可剪接至外显子2,外显子2含有蛋白质编码序列的起始位点。致病性六核苷酸重复扩增定位在外显子1A与外显子1B之间。因此,从外显子1A启动的转录可产生含重复的RNA,而从外显子1B启动则不产生。

[0918] 为了将C9orf72重复扩增疾病在小鼠中建模,如下所描述,将等位基因系列构建于小鼠胚胎干(ES)细胞中,其中,将来自人类C9orf72基因的片段(包括外显子1A的一部分、位于1A与1B之间的内含子、外显子1B的全部以及下游内含子的一部分)精确地放置在小鼠C9orf72基因中一个等位基因的其同源位置处。参见,例如,US 2018/0094267及WO 2018/064600,其各自通过援引而出于全部目的整体并入本文。将一系列六核苷酸重复扩增放置在于人类基因中发现的位置处,该位置从正常的三次重复到高达致病性的600次重复。

[0919] 在培养基中使携带不同重复扩增的小鼠ES细胞克隆分化为运动神经元,以研究该扩增对与ALS相关的细胞类型的效应。在检查从经基因修饰的人源化C9orf72等位基因产生的转录本时发现,存在从三次重复的正常对照中占主导地位的外显子1B剪接转录本到具有更长重复扩增的等位基因中外显子1A剪接转录本的出现增加的转变。还观察到其丰度直接与六核苷酸重复扩增长度相关的含未剪接内含子的转录本的累积,表明自我前馈环(selfish feed-forward loop)中的重复扩增越长,则自C9orf72基因产生的含重复的转录本越多。靶向含重复序列的内含子转录本以摧毁或激活二肽重复蛋白质合成的模板,同时保留正常C9orf72mRNA及蛋白质的合成,有望成为针对C9orf72重复扩增疾病的安全且有效的治疗策略。

[0920] 一种用以降低含C9orf72重复序列的RNA的可能途径是通过RNA干扰的天然过程实现,在该过程中,siRNA引导靶标RNA被RNA诱导沉默复合物直接裂解,之后RNA裂解片段被细胞核酸酶降解。然而,RNA干扰主要是细胞质过程,预期其不会作用于保留在细胞核中的RNA。含内含子的RNA通常是短寿命的,或作为mRNA前体,其被迅速剪接为成熟mRNA,或作为被剔除的内含子,在细胞核中迅速降解。因此,可合理地预期,含内含子的RNA将不能用于通

过RNA干扰进行靶向。

[0921] 然而,已经证明,靶向与GGGGCC重复扩增相邻的内含子序列的siRNA促进含内含子的C9orf72 RNA的累积的降低,同时对C9orf72成熟RNA具有极低效应或无效应。靶向内含子的siRNA也降低二肽重复蛋白质的产生。这些出乎意料的实验结果指示,细胞中累积的具有C9orf72六核苷酸重复扩增的含内含子的RNA易受RNA干扰的影响。结果显示,负责二肽重复蛋白质合成的相当一部分含内含子的C9orf72RNA位于细胞质中。相比之下,靶向C9orf72 mRNA蛋白质编码序列的siRNA产生mRNA的强烈敲低,但对于含内含子的转录本无效,并且未明显降低二肽重复蛋白质合成。靶向内含子的siRNA与靶向mRNA的siRNA之间的结果发散性表明,两类靶向序列存在于未共价连接的独立RNA上。

[0922] 本文所述的方法及组合物提供二肽重复蛋白质合成的治疗性降低,该蛋白质是C9orf72重复扩增疾病的主要致病组分,同时隔离C9orf72mRNA,从而避免C9orf72蛋白质降低的可能的副作用,而该可能的副作用在使用治疗性策略诸如使用靶向细胞核中的初代C9orf72转录本的反义寡核苷酸时将会发生。

[0923] II.C9orf72、六核苷酸重复扩增序列、及相关疾病

[0924] 肌萎缩侧索硬化(ALS),也称为Lou Gehrig氏病,是成年发作频率最高的麻痹疾患,以上和/或下运动神经元的损失为特征。美国每年有多达20,000人出现ALS,其中约5,000例为新病例。额颞叶痴呆(FTD;也称为Pick氏病、额颞叶退行或FTLD)是一组疾患,由脑的额叶或颞叶中的进行性细胞退行造成。据报导,FTD占全部痴呆病例的10%至15%。位于(且任选跨越)人类C9orf72基因的两个非编码外显子即外显子1A与1B之间的六核苷酸重复扩增序列已经与ALS及FTD两者相关联。据估计,GGGGCC(SEQ ID NO:1)六核苷酸重复扩增占家族性ALS病例的约50%以及许多非家族性ALS病例。因为C9orf72重复序列扩增是显性突变,其基本上总是作为杂合突变而发生。其存在于约25%的家族性FTD病例中及约8%的散发性FTD病例中。

[0925] 已经报导了许多与C9orf72中六核苷酸重复扩增序列相关的病理学方面,例如,RNA团簇的重复序列长度依赖性形成、特异性RNA结合蛋白质的隔离、及其神经元中重复序列相关非AUG(AUG)翻译所致的二肽重复蛋白质(例如,聚(甘氨酸-丙氨酸)、聚(甘氨酸-脯氨酸)、聚(甘氨酸-精氨酸)、聚(丙氨酸-脯氨酸)及聚(脯氨酸-精氨酸))的累积及集聚。

[0926] 尽管已经报导了C9orf72调控胞吞运输,但C9orf72的细胞功能大多仍是未知。事实上,C9orf72是编码其功能未知的未表征蛋白质的基因。

[0927] 已经报导了小鼠C9orf72转录本变体。参见,例如,Koppers et al.(2015)Ann.Neurol.78:426-438及Atkinson et al.(2015)Acta Neuropathologica Communications 3:59,其各自通过援引而出于全部目的整体并入本文。三种所报导的小鼠C9orf72转录本变体的基因组信息也可于Ensembl网站在登记号ENSMUST00000108127(V1)、ENSMUST00000108126(V2)及ENSMUST00000084724(V3)下获得。示例性非人(例如,啮齿动物)C9orf72 mRNA及氨基酸序列详述于SEQ ID NOS:40至43、123及124中。小鼠C9orf72的mRNA及氨基酸的序列可分别见于GenBank登录号NM_001081343及NP_001074812,并且通过援引而出于全部目的整体并入本文。NM_001081343.1及NP_001074812.1的序列分别于SEQ ID NOS:40(及SEQ ID NO:123)及41中详述。大鼠C9orf72的mRNA及氨基酸的序列可分别见于GenBank登录号NM_001007702及NP_001007703,并且通过援引而出于全部目的整体并入

本文。NM_001007702.1及NP_001007703.1的序列分别于SEQ ID NOS:42(及SEQ ID NO:124)及43中详述。

[0928] 人类C9orf72转录本变体也是已知的。一种人类C9orf72转录本变体缺乏中心及3'编码区域中的多个外显子,并且其3'端外显子延伸超过变体3(见下文)中使用的剪接位点,其导致与变体3相比新颖的3'未翻译区域(UTR)。该变体编码显著较短的多肽,并且其C端氨基酸与由两种其他变体编码者相比截然不同。该变体的mRNA及氨基酸的序列可分别见于GenBank登录号NM_145005.6及NP_659442.2,并且通过援引而出于全部目的整体并入本文。NM_145005.6及NP_659442.2的序列分别于SEQ ID NOS:94及SEQ ID NO:110中详述。与变体3相比,第二人类C9orf72转录本变体(2)在5'未翻译区域(UTR)有所不同。该变体的mRNA及氨基酸的序列可分别见于GenBank登录号NM_018325.4及NP_060795.1,并且通过援引而出于全部目的整体并入本文。NM_018325.4及NP_060795.1的序列分别于SEQ ID NOS:26及SEQ ID NO:27中详述。第三人类C9orf72转录本变体(3)含有三种所报导变体中最长的序列并且编码较长的同种型。该变体的mRNA及氨基酸的序列可分别见于GenBank登录号NM_001256054.2及NP_001242983.1,并且通过援引而出于全部目的整体并入本文。NM_001256054.2及NP_001242983.1的序列分别于SEQ ID NOS:28(及SEQ ID NO:121)及SEQ ID NO:29中详述。变体2及3编码相同蛋白质。

[0929] III.在C9orf72基因座处含有六核苷酸重复扩增的ES细胞克隆的生成

[0930] 如上所述,肌萎缩侧索硬化(ALS)是进行性神经退行性疾患,其造成运动神经元死亡,导致麻痹。在美国,每年有五千人被诊断患有ALS。该疾病的10%是在家族中传播,并且这一类的最常见病因是C9orf72基因座处的GGGGCC(SEQ ID NO:1)六核苷酸重复扩增。健康个体典型具有少于30个重复序列,而受影响的患者一般具有该重复序列的超过1000个拷贝。因为缺乏可用的动物模型,重复序列的这一长拉伸如何造成疾病的精确机制尚未阐明。六核苷酸重复序列中的高GC含量使得其难以合成具有该重复序列的DNA片段并且难以在微生物中维持该重复序列。因此,制备诸如靶向载体的材料以生成转基因动物非常具有挑战性。这是为何迄今为止可得到极少可用于C9orf72重复扩增类型的ALS动物模型的主要原因之一,并且本领域中存在对更优选动物模型的强烈需求。

[0931] 为绕过这些用于靶向载体构建的困难的中间步骤,采取了扩增已经被插入基因组中正确位置中的相对较短的六核苷酸重复序列的途径,而非进行靶向载体及ES细胞靶向的从头产生。作为用于重复扩增的起始材料,使用了杂合ES细胞克隆,该细胞克隆中,小鼠C9orf72基因座的一部分被替换为含有GGGGCC(SEQ ID NO:1)六核苷酸的92x重复的人类对应物。参见图1。先前研发了这一等位基因。参见,US 2018/0094267及WO 2018/064600,其各自通过援引而出于全部目的整体并入本文。

[0932] 当DNA双链断裂(DSB)出现于哺乳动物细胞染色体中时,该细胞通过同源重组(HR)或异源重组途径诸如非同源末端连接(NHEJ)或单链退火(SSA)来识别该DSB并修复该损害。在HR中,受损的染色单体利用非受损的染色单体作为其修复的模板。在这一途径中,一旦DSB出现,即通过包括3'链截断在内的多种过程修饰DSB。所暴露的5'单链通过多个分子的Rad51包被,然后,Rad51丝状构建体开始搜索该基因组中的同源序列。Rad51丝状构建体能够将其自身插入双链DNA中,在插入之处,所暴露的5'单链具有同源性。所侵入的5'链使用该Rad51丝状构建体进入之处的序列启动恢复。Rad51丝状构建体用以找到同源序列的准确

度对于保持正确的DNA序列是关键。用于扩增重复序列的途径是基于以下假定:通过将DSB引入邻近六核苷酸重复序列处,使得所暴露的5'链仅含有重复性序列,Rad51丝状构建体可能进入错误位置,并且其可能误导通过Rad51丝状构建体进行的同源性搜索,导致重复序列的扩增或收缩。参见图2。

[0933] 具体而言,本文所用的途径是通过使用CRISPR/Cas9核酸酶将DSB引入邻近六核苷酸重复序列终止位点处,在人源化小鼠胚胎干细胞中扩增C9orf72内含子1处的现有GGGGCC(SEQ ID NO:1)重复序列。在尽可能邻近六核苷酸重复序列处设计若干个导引RNA(gRNA),以在截断之后暴露5'单链处的重复性序列。参见图3及表4。

[0934] 表4.C9orf72 gRNA靶标序列.

导引 RNA	靶标序列 (NGG PAM 的上游)	SEQ ID NO
5' gRNA-1	GCTATGCGATCGCCGTCTCG	30
5' gRNA-2	CCCCGGCCCCGGCCCCGAGA	31
5' gRNA-3	CGCCGTCTCGGGGCCGGGGC	32
5' gRNA-4	CCGTCTCGGGGCCGGGGCCG	33
3' gRNA-1	CGGCCGGCCCTCGAGGGTCT	34
3' gRNA-2	GGCCGGGGCCGAGACCCTCG	35
3' gRNA-3	GAGGGTCTCGGCCCCGGCCC	36
3' gRNA-4	CTCGGCCCCGGCCCCGGCCC	37
导引 RNA	导引序列	SEQ ID NO
5' gRNA-1	GCUAUGCGAUCGCCGUCUCG	84
5' gRNA-2	CCCCGGCCCCGGCCCCGAGA	85
5' gRNA-3	CGCCGUCUCGGGGCCGGGGC	86
5' gRNA-4	CCGUCUCGGGGCCGGGGCCG	87
3' gRNA-1	CGGCCGGCCCUCGAGGGUCU	88
3' gRNA-2	GGCCGGGGCCGAGACCCUCG	89
3' gRNA-3	GAGGGUCUCGGCCCCGGCCC	90
3' gRNA-4	CUCGGCCCCGGCCCCGGCCC	91

[0935]

导引 RNA	crRNA 序列	SEQ ID NO
5' gRNA-1	GCUAUGCGAUCGCCGUCUCGGUUUUAGAGCUAUGCUGUUUUG	56
5' gRNA-2	CCCCGGCCCCGGCCCCGAGAGUUUUAGAGCUAUGCUGUUUUG	57
5' gRNA-3	CGCCGUCUCGGGGCCGGGGCGUUUUAGAGCUAUGCUGUUUUG	76
5' gRNA-4	CCGUCUCGGGGCCGGGGCCGGUUUUAGAGCUAUGCUGUUUUG	77
3' gRNA-1	CGGCCGGCCCUCGAGGGUCUGUUUUAGAGCUAUGCUGUUUUG	78
3' gRNA-2	GGCCGGGGCCGAGACCCUCGGUUUUAGAGCUAUGCUGUUUUG	79
3' gRNA-3	GAGGGUCUCGGCCCCGGCCCGUUUUAGAGCUAUGCUGUUUUG	80
3' gRNA-4	CUCGGCCCCGGCCCCGGCCCGUUUUAGAGCUAUGCUGUUUUG	81
导引 RNA	TracrRNA 序列	SEQ ID NO
全部	GUUGGAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU	82

[0936] 目标为获得gRNA,其裂解尽可能接近重复并且具有高裂解效率。使用含有与含人源化的92x重复序列的等位基因相同的序列的质体(8028Stvec),于体外测试DNA裂解效率。从重复的5'侧及3'侧两侧选择具有最高裂解效率的gRNA。这些是5'gRNA-1及3'gRNA-2。参见图4。

[0937] 为了在ES细胞中修饰DNA，通过电穿孔将SpCas9(酿脓链球菌Cas9;Thermo Fisher)及gRNA引入含C9orf72人源化的92x重复的ES细胞(MAID8029a)中作为RNA与蛋白质复合体(RNP)。选取ES细胞克隆，该克隆于96孔培养板中生长，并且将基因组DNA纯化以用于分析。进行了三个实验：(1)将单DSB引入重复的5'侧处；(2)将单DSB引入重复的3'侧处；以及(3)将DSB引入重复的5'侧及3'侧两处。参见图5。使用常规双引物PCR(图6A)及使用三个引物的使用AmplideX PCR/CE C9orf72试剂盒的引物PCR(图6B)进行该分析。对于引物PCR，使用来自ES细胞的基因组DNA作为模板。使用三个引物：定位于重复序列的5'外侧的引物、定位于重复序列的3'外侧的引物、以及退火至重复序列内部的重复性序列的第三引物，如图6B中所示。这一PCR反应产生众多不同尺寸的PCR产物，如图6B中所示。因为第三引物可于重复序列的任意给定单元中引发聚合酶反应，所获得的PCR产物的数目基本上是克隆所具有的重复数目。将这些PCR产物运行毛细管电泳，检测其萦光信号(引物之一经萦光标记)。

[0938] 根据制造商的使用说明，使用AmplideX PCR/CE C9orf72试剂盒(Asuragen)来证实内源性C9orf72 ES细胞克隆中如SEQ ID NO:1详述的六核苷酸序列的实例数目。使用来自3x重复序列克隆、92x重复序列克隆及30x重复序列克隆的经纯化的mESC基因组总DNA作为对照。使用表5中的引物及来自AmplideX PCR/CE C9orf72试剂盒的重复序列特异性引物的PCR是在ABI 9700热循环仪(Thermo Fisher)上执行。扩增子通过毛细管电泳于ABI 3500xL GeneScan上使用POP-7聚合物(Thermo Fisher)及NuSieve琼脂糖凝胶(Lonza)测定尺寸。将2-log DNA阶梯(New England BioLabs)分子量标记加载到琼脂糖凝胶进行比较，使用SYBR金核酸染色(Thermo Fisher)将条带可视化。

[0939] 表5.用于PCR的引物.

[0940]

引物名称	序列	SEQ ID NO:
2-正向引物	TGCGCCTCCGCCGCCGCGGGCGCAGGCACCGCAACCGCA	92
2-反向引物	CGCAGCCTGTAGCAAGCTCTGGAACTCAGGAGTCG	93

[0941] 将DSB引入重复序列的任一侧导致重复序列处的尺寸改变，如使用常规双引物PCR所证明。参见图7A及图7B。每种条件下，88种克隆中的16种的结果显示于图7A及图7B中。选择具有最大扩增的5'单DSB克隆及3'单DSB克隆进行进一步分析，使用引物PCR来证实重复扩增。引物PCR证实，克隆9253D-B2(5'单DSB)具有145次重复，而克隆9253A-B1(3'单DSB)具有130次重复及92次重复。参见图8及图9。图8显示毛细管电泳的结果。信号强度示于Y轴上，而PCR产物尺寸示于X轴上。读出是峰的数目。上图来自92x亲本克隆，并且存在92个峰。在中图中计数出150个峰，下图中计数出130x峰。在下图中观察到两个高峰。这可能是因为选取了ES细胞克隆而非纯合克隆。

[0942] 重复相同的实验，获得类似的结果。参见图10。对于每一条件(单5'DSB、单3'DSB、或5'DSB及3'DSB两者)，测试了88个克隆。重复序列处的DSB造成重复序列不稳定。重复序列的一侧或两侧处的DSB导致重复序列扩增(从92x到多达约300x)。对不同克隆中扩增或收缩的完全分析提供于表6中。通过TAQMAN测定或Asuragen AmplideX PCR/CE C9orf72试剂盒调用完全收缩。

[0943] 表6.第一次重复扩增的重复扩增或缩短的评估.

[0944]

重排	5’ & 3’ DSB (%)	Single 5’ DSB (%)	Single 3’ DSB (%)
扩增 (>750bp = 92x)	4.5	18.2	5.6
部分缩短 (<750bp, >200bp)	9.1	9.1	9.1
保持 (750bp)	19.3	36.4(51.1)*	5.6(28.4)*
完全缩短 (<200bp)	46.6	21.6	55.8
未侦测	20.5	14.8	27.3
总计	100	100	100

[0945] *观察到多个条带

[0946] 通过一起引入5’DSB及3’DSB两者而获得最大重复序列。通过引入单DSB获得的最大重复为大约250X，其中，3’gRNA位点仍完整且该重复序列可扩增。为了进一步扩增该重复序列，使用250x重复克隆(MAID9253D-C5)，并通过引入如上所述的Cas9/gRNA RNP而将DSB引入到3’末端。使用250x等位基因，将该重复进一步扩增至大约600x的尺寸。参见图11A、图11B及表7。

[0947] 表7.利用单一3’DSB的第二次重复扩增后的重复扩增或缩短的评估

[0948]

亲代克隆	8029,92x(%)	9253D-C5,250x(%)
扩增(>亲代克隆)	5.6	13.6
部分缩短(<亲代克隆,>200bp)	9.1	43.2
保持(=亲代克隆)	5.6(28.4)*	30.7(43.2)*
完全缩短(<200bp)	55.8	4.5
未检测	27.3	2.3
总计	100	100

[0949] *观察到多个条带

[0950] 实施例2:在C9orf72基因座处含有六核苷酸重复扩增的运动神经元或脑组织的分析

[0951] 如上所描述，使用AmplideX PCR/CE C9orf72试剂盒(Asuragen)证实了小鼠中六核苷酸重复的尺寸的稳定性。

[0952] 如US 2018/0094267及WO 2018/064600中所描述，检查了包含野生型C9orf72基因座(对照)或来自实施例1的经基因修饰的C9orf72基因座的源自小鼠胚胎干细胞的运动神经元(ESMN)中的RNA转录本，两篇专利各自通过援引将其全部内容并入本文。在包含野生型C9orf72基因座(对照)或来自实施例1的经基因修饰的C9orf72基因座的源自亲本胚胎干细胞的ESMN中，评估了RNA团簇及二肽重复蛋白质水平。材料及方法描述如下。

[0953] 此外，如US 2018/0094267及WO 2018/064600中所描述，检查了包含野生型C9orf72基因座(对照)或来自实施例1的经基因修饰的C9orf72基因座的小鼠脑组织及亲本胚胎干细胞中的RNA转录本，两篇专利各自通过援引将其全部内容并入本文。在包含野生型C9orf72基因座(对照)或来自实施例1及2的经基因修饰的C9orf72基因座的源自亲本胚胎干细胞的ESMN中，评估了RNA团簇及二肽重复蛋白质水平。材料及方法描述如下。

[0954] 源自胚胎干细胞的运动神经元

[0955] 将来自实施例1的胚胎干细胞(ESC)于胚胎干细胞培养基(ESM;DMEM+15％胎牛血清+青霉素/链霉素+谷氨酰胺+非必需氨基酸+核苷+β-巯基乙醇+丙酮酸钠+LIF)中培养2天，其间每天更换培养基。在胰蛋白酶化的前1小时，将ES培养基替换为7mL的ADFNK培养基

(高级DMEM/F12+神经基底培养基+10％敲除血清+青霉素/链霉素+谷氨酰胺+β-巯基乙醇)。抽吸ADFNK培养基,并使用0.05％胰蛋白酶-EDTA将ESC胰蛋白酶化。将成球丸的细胞再悬浮于12mL的ADFNK中,并于悬浮于中生长两天。将细胞于以视黄酸(RA)、平滑激动剂(smoothened agonist)以及嘌吗啡胺(purmorphamine)补充的ADFNK中进一步培养4天,以获得类四肢运动神经元(ESMN)。将解离的运动神经元铺板,并于源自胚胎干细胞的运动神经元培养基(ESMN;神经基底培养基+2％马血清+B27+谷氨酰胺+青霉素/链霉素+β-巯基乙醇+10ng/mL GDNF、BDNF、CNTF)中培养熟化。

[0956] 定量聚合酶链式反应

[0957] 提取来自每一样本的总RNA,并且使用位于多个区域侧翼的引物及检测经修饰的C9orf72基因座的那些区域的探针进行逆转录。可检测的区域包括那些跨越小鼠序列与人类序列连接者、仅人类序列或仅小鼠序列。GAPDH、DROSHA或β2-微球蛋白的qPCR是使用即用型试剂盒的探针及引物执行的。

[0958] 具体而言,RNA是自包含野生型(WT)C9orf72基因座(对照)或经基因修饰的C9orf72基因座的源自胚胎干细胞的运动神经元(ESMN)分离。在其他试验中,RNA也自包含野生型(WT)C9orf72基因座(对照)或经基因修饰的C9orf72基因座的分离自小鼠的亲本胚胎干(ES)细胞或总脑分离。

[0959] 总RNA使用Direct-zol RNA Miniprep plus试剂盒根据制造商的方案(Zymo Research)分离。使用Turbo无DNA试剂盒,根据制造商的方案(Invitrogen),用DNase处理总RNA,并稀释至20ng/μL。使用Quantitect Probe RT-PCR试剂盒(Qiagen),于一步反应中执行逆转录(RT)及PCR。qRT-PCR反应含有2μL RNA及8μL混合物,该混合物含有RT-PCR预混液、ROX染剂、RT混合液及20X基因特异性引物-探针混合液,制成最终体积为10μL。

[0960] 除非另做标注,否则最终的引物及探针浓度分别为0.5μM及0.25μM。qRT-PCR是于ViiA™7实时PCR检测系统(ThermoFisher)上执行的。PCR反应在光学384孔板中平行进行四次,其中,RT步骤在45℃进行10min,之后在95℃进行10min,以及在95℃ 5s、60℃ 30s的两步循环,进行45次循环。每一分析中使用的引物及探针的序列以及SEQ ID NO(A、B、C、D、E、F、G、H、I及J)提供于表8中。

[0961] 表8.引物及探针

[0962]

分析 A	
正向引物	CATCCCAATTGCCCTTTCC (SEQ ID NO: 61)
反向引物	CCCACACCTGCTCTTGCTAGA (SEQ ID NO: 62)
探针	TCTAGGTGGAAAGTGGG (SEQ ID NO: 63)
分析 B	
正向引物	GAGCAGGTGTGGGTTTAGGA (SEQ ID NO: 64)
反向引物	CCAGGTCTCACTGCATTCCA (SEQ ID NO: 65)
探针	ATTGCAAGCGTTCGGATAATGTGAGA (SEQ ID NO: 66)
分析 C	
正向引物	GATAGTCGACATCCCTGCATC (SEQ ID NO: 95)
反向引物	GGTGGCGAGTGGCTATTG (SEQ ID NO: 96)
探针	AAGCGTTCGGATAATGTGAGACCTGG (SEQ ID NO:

[0963]

	97)
分析 D	
正向引物	GCTGTCACGAAGGCTTTCTTC (SEQ ID NO: 67)
反向引物	GCACTGCTGCCAACTACAAC(SEQ ID NO: 68)
探针	TCAATGCCATCAGCTCACACCTGC(SEQ ID NO: 69)
分析 E	
正向引物	TCTCACAGTACTCGCTGAGGGTGA (SEQ ID NO: 98)
反向引物	AAGAGCAGGTGTGGGTTTAG (SEQ ID NO: 99)
探针	CGGTTGTTTCCCTCCTTGT (SEQ ID NO: 100)
分析 F	
正向引物	CCCACTACTTGCTCTCACAG (SEQ ID NO: 101)
反向引物	TACAGGCTGCGGTTGTTT (SEQ ID NO: 102)
探针	ACTCGCTGAGGGTGAACAAGAAA (SEQ ID NO: 103)
分析 G	
正向引物	AAGAGGCGCGGGTAGAA (SEQ ID NO: 70)
反向引物	CAGCTTCGGTCAGAGAAATGAG (SEQ ID NO: 71)
探针	CTCTCCTCAGAGCTCGACGCATTT (SEQ ID NO: 72)
分析 H	
正向引物	CTGCACAATTTCAGCCCAAG (SEQ ID NO: 73)
反向引物	CAGGTCATGTCCCACAGAAT (SEQ ID NO: 74)
探针	CATATGAGGGCAGCAATGCAAGTC (SEQ ID NO: 75)
分析 I	
正向引物	CGAGTGGGTGAGTGAGGA (SEQ ID NO: 104)
反向引物	TTCTACCCGCGCCTCTT (SEQ ID NO: 105)
探针	ATCCTGGCGGGTGGCTGTTT (SEQ ID NO: 106)
分析 J	
正向引物	CGGATAATGTGAGACCTGGAAT (SEQ ID NO: 107)
反向引物	AAAGGTAGCCGCCAACAA (SEQ ID NO: 108)
探针	ACCATCTCCTGCTGTTGCCAAGA (SEQ ID NO: 109)

[0964] Western Blot分析

[0965] 收集经分化的胚状体(EB),并于SDS样本缓冲液(2%SDS、10%甘油、5%β-巯基乙醇、60mM TrisHCl,pH 6.8、溴酚蓝)中均质化。蛋白质提取物使用RC DC蛋白质测定(BioRad)定量。提取物(10μg)于4-20%SDS-PAGE凝胶(ThermoFisher)上运行,使用iBLOT转移单元(ThermoFisher)转移到硝基纤维素膜。免疫印迹使用抗C9orf72及GAPDH的一级抗体(Millipore)探查。通过如下检测经结合的抗体:使用缀合至辣根过氧化物酶的二级抗体(Abcam)温育,之后使用SuperSignal West Pico化学发光底物(Thermo Scientific)进行化学发光。使用Full Speed Blue敏感性医学X-射线膜(Ewen Parker XRay Corporation)通过自动放射照相术检测信号。使用ImageJ计算相对蛋白质水平。数据未显示。

[0966] 用于检测正义或反义RNA转录产物的荧光原位杂交(FISH)

[0967] 使用荧光原位杂交(FISH)测定从SEQ ID NO:1的六核苷酸重复序列转录的RNA在如上文所述生成的源自胚胎干细胞的运动神经元(ESMN)中的定位。简而言之,使ESMN在四孔室载玻片(Lab-Tek II区室载玻片系统,ThermoFisher Scientific)中生长,并使用4%PFA(Electron Microscopy Sciences)于PBS中的溶液固定。然后,使用焦碳酸二乙酯(DEPC)PBS/0.2%Triton X-100(Fisher Scientific,目录号BP151)将细胞透化,以DEPC-PBS洗涤,封闭,用LNA寡核苷酸染色以进行对RNA转录产物的检测,如下文所描述。染色之后,载玻片接着用适宜的荧光染剂温育,用Fluoromount G(Southern Biotech)封固,并使用共焦显微镜可视化。

[0968] 对于LNA探针，载玻片于66℃，用由50％甲酰胺(IBI Scientific，目录号IB72020)、DEPC 2×SSC[300mM氯化钠、30mM柠檬酸钠(pH7.0)]、10％(w/v)硫酸葡聚糖(Sigma-Aldrich，目录号D8960)及DEPC 50mM磷酸钠(pH 7.0)所组成的缓冲剂预杂交30分钟。然后将杂交缓冲液吸取排除，向每一载玻片添加400μL的于杂交缓冲液中的40nM LNA探针混合液，并且于黑暗中在66℃温育3小时(对于LNA探针)。经LNA探针温育的载玻片在室温于DEPC 2×SSC/0.1％吐温20(Fisher Scientific，目录号BP337)中冲洗一次，在65℃于DEPC 0.1×SSC中冲洗三次。接着，载玻片用1μg/mL DAPI(Molecular Probes Inc.)温育。

[0969] 在另一个实验中，载玻片用由50％甲酰胺(IBI Scientific，目录号IB72020)、DEPC 2×SSC[300mM氯化钠、30mM柠檬酸钠(pH 7.0)]、10％(w/v)硫酸葡聚糖(Sigma-Aldrich，目录号D8960)及DEPC 50mM磷酸钠(pH 7.0)所组成的缓冲剂于66℃(对于LNA探针)或于55℃(对于DNA探针)预杂交30分钟。然后将杂交缓冲液吸取排除，向每一载玻片添加400μL的于杂交缓冲液中的200ng/mL DNA探针混合液，并且于黑暗中在55℃温育3小时。载玻片用DNA探针温育，用40％甲酰胺于2×SSC中洗涤三次，并于PBS中简单洗涤一次。接着，载玻片用1μg/mL DAPI(Molecular Probes Inc.)温育。

[0970] 本实施例中使用的LNA及DNA寡核苷酸探针的序列及SEQ ID NO以及这些探针的杂交条件提供于下表9中。锁核酸(LNA)是核酸类似物，其中，核糖部分经连接2’氧与4’碳的外桥修饰。

[0971] 表9.LNA及DNA探针.

[0972]

探针	序列 (SEQ ID NO)	杂交方法

[0973]

LNA 正义 G_4C_2 RNA	TYE563-CCCCGGCCCCGGCCCC (SEQ ID NO: 38)	66°C 杂交及于 0.1 X SSC 中洗涤
LNA 反义 G_4C_2 RNA	TYE563-GGGGCCGGGGCCGGGGGG CCCC (SEQ ID NO: 39)	66°C 杂交及于 0.1 X SSC 中洗涤
DNA 正义 G_4C_2 RNA	CCCCGGCCCCGGCCCCGG—Cy3 (SEQ ID NO: 44)	55°C 杂交及于 2 X SSC 中洗涤
DNA 反义 G_4C_2 RNA	GGGGCCGGGGCCGGGGC-Cy3 (SEQ ID NO: 45)	55°C 杂交及于 2 X SSC 中洗涤

[0974] 二肽重复蛋白质产物的检测(免疫萦光及western slot blot)

[0975] 使用免疫萦光来评估如上所述的源自胚胎干细胞的运动神经元(ESMN)中二肽重复蛋白质产生。简而言之，使ESMN于四孔室载玻片(Lab-Tek II区室载玻片系统，ThermoFisher Scientific)中生长，并使用4％PFA(Electron Microscopy Sciences)于PBS中的溶液固定。然后，使用焦碳酸二乙酯(DEPC)PBS/0.2％Triton X-100(Fisher Scientific，目录号BP151)将细胞透化，以DEPC-PBS洗涤，封闭，用抗polyGA抗体染色以进行对RAN转录产物的检测，如下文所描述。染色之后，载玻片接着用适宜的萦光染剂温育，用Fluoromount G(Southern Biotech)封固，并使用共焦显微镜可视化。

[0976] 透化之后，载玻片用稀释于Tris缓冲盐水(pH 7.4)中的5％正常驴血清及0.2％Triton X100(TBS-T)封闭。载玻片用稀释于TBS-T中的针对poly-GA的一级抗体(Millipore)及5％正常驴血清在4℃温育越夜。用TBS-T洗涤3次之后，在室温，载玻片用偶联至Alexa 488或555的物种特异性二级抗体(1:1000，于TBS-T中，ThermoFisher)及DAPI(1μg/mL)(Molecular Probes Inc.)温育1小时。用TBS-T洗涤3次之后，载玻片用Fluoromount

G(Southern Biotech)封固,并且使用共焦显微镜可视化。

[0977] 对于狭缝印迹(slot blot)测定,收集经分化的胚状体(EB),并于SDS样本缓冲液(2%SDS、10%甘油、5%β-巯基乙醇、60mM TrisHCl,pH 6.8、溴酚蓝)中均质化。蛋白质提取物使用RC DC蛋白质测定(BioRad)定量。在真空下,使用Bio-Slot 48空微滤系统(Bio-Rad)将含有0μg、1.25μg、2.5μg、5μg、10μg或20μg的裂解液固定于硝基纤维素膜。该膜于TBS-T中洗涤,并使用针对poly(GP)的抗体(1:5,000,Novus biologicals)及针对poly GA的抗体(1:5000,Millipore)进行印迹。将膜用HRP缀合的二级抗体温育之后,通过ECL和Western Blotting Detection System检测系统(Pierce)将条带可视化。

[0978] 结果

[0979] 测试了源自胚胎干细胞的运动神经元(ESMN),其包含人源化C9orf72等位基因的等位基因系列,该等位基因系列具有G_4C_2六核苷酸序列的3次重复、92次重复、250次重复、300次重复、500次重复或600次重复。如图12C、图12D、图12G及图12H所示,包含C9orf72处的六核苷酸重复扩增序列的ESMN显示,保留内含子1序列的C9orf72mRNA转录本的表达增加。此外,较大的G_4C_2重复扩增对外显子1A的使用增加而对外显子1B的使用降低。分别参见图12A及图12B,以及分别参见图12E及图12F。含有G_4C_2重复扩增的ESMN也含有细胞核及细胞质正义及反义C9orf72 RNA团簇(数据未显示)。正义G_4C_2团簇较大且分布于核仁中(数据未显示)。具有大于300次重复序列扩增的RNA团簇在ESMN核仁中的定位再现了源自ALS患者的细胞中的病理学发现。此外,G_4C_2六核苷酸序列重复次数的增加与从六核苷酸重复转录本翻译(通过RAN翻译,一种非AUG机制)的二肽重复蛋白质(polyGA和polyGP)的增加直接相关。参见,例如,图13。总之,源自等位基因序列ES细胞的神经元再现ALS疾病的分子标志(正义及翻译重复RNA团簇、定位在核仁处的重复序列RNA团簇、五种形式的二肽重复蛋白质中的至少两者、以及含内含子转录本的累积增加),支持使用本文所公开的非人类动物模型作为神经退行性疾病的疾病模型。

[0980] 在两种类型的ESMN中重复用于表8中的测定C、B、F、H、E及D的定量PCR反应:类轴下运动神经元(MN)及类四肢运动神经元(MN)。在模型中,类轴下MN支配轴下MN。由轴下MN支配的肌肉的示例包括肋间肌、隔膜及腹壁肌肉。类四肢MN是那些起源于横向运动柱并且支配远程四肢肌肉,诸如位于前肢及后肢中的那些(胫骨前肌、腓肠肌及臀部肌群)。轴下MN主要通过添加视黄酸(RA)及声波刺猬激动剂(SAG)的方案产生。类四肢MN可通过将1μM噻吗啡胺添加到RA及SAG顶部而生成,可借此生成类四肢MN。

[0981] 测试了运动神经元,其包含人源化C9orf72等位基因的等位基因系列,该等位基因系列具有G_4C_2六核苷酸序列的3次重复、92次重复、300次重复、500次重复或600次重复。如图15A及图15B中所示,较大的G_4C_2重复序列扩增对外显子1A的使用增加而对外显子1B的使用降低。如图15C及图15D所示,包含C9orf72处的六核苷酸重复扩增序列的ESMN显示,保留内含子1序列的C9orf72 mRNA转录本的表达增加。如图15E中所示,较大的G_4C_2重复序列扩增具有增加的经剪接前体的表达。如图15F中所示,经剪接C9orf72 mRNA的表达并未随着重复尺寸变化而改变很多。

[0982] 实施例3:在C9orf72基因座处含有六核苷酸重复扩增的小鼠的生成

[0983] 使用VELOCIMOUSE®方法,使用实施例1中所述的300x重复ES细胞克隆生成F0小鼠。参见,例如,US 7,576,259;US 7,659,442;US 7,294,754;US 2008/0078000;以及

Poueymirou et al.(2007)Nat.Biotechnol.25(1):91-99,其各自通过援引而出于全部目的整体并入本文。在VELOCIMOUSE®方法中,通过激光辅助注射将所靶向的小鼠胚胎干(ES)细胞注射到前桑椹胚阶段的胚胎例如8细胞阶段的胚胎中,这有效地导致完全源自ES细胞的F0小鼠生成。

[0984] 在实施例3中,测试了来自C9orf72 300x重复序列小鼠的脑干及脊髓样本的RNA转录本(即,人源化C9orf72等位基因,其具有G_4C_2六核苷酸序列的300次重复)及来自对照C9orf72 3x重复序列小鼠的脑干及脊髓样本的RNA转录本(即,人源化C9orf72等位基因,其具有G_4C_2六核苷酸序列的3次重复)。如实施例3中所述,评估了来自C9orf72 300x重复序列小鼠的脑干及脊髓样本内的RNA团簇及二肽重复蛋白质水平。如图14B中所示,来自C9orf72 300x重复序列小鼠的脑干及脊髓样本显示,保留内含子1序列的C9orf72 mRNA转录本的表达增加。此外,还测试了脊髓样本(即,L4/L5腰部脊髓运动神经元),并且其显示含有细胞核及细胞质正义及反义C9orf72 RNA团簇(数据未显示)。同样,这些样本中从六核苷酸重复序列转录本翻译(通过RAN翻译,一种非AUG机制)的二肽重复蛋白质(polyGA)的存在增加(数据未显示)。

[0985] 类似地,测试了源自胚胎干细胞的运动神经元(ESMN),其包含人源化C9orf72等位基因的等位基因系列,该等位基因系列具有G_4C_2六核苷酸序列的3次重复、92次重复、00次重复、500次重复或600次重复。如图14A所示,包含C9orf72处的六核苷酸重复扩增序列的ESMN显示,保留内含子1序列的C9orf72 mRNA转录本的表达增加。含有G_4C_2重复扩增的ESMN还含有细胞核及细胞质正义及反义C9orf72 RNA团簇,并且从六核苷酸重复序列转录本翻译(通过RAN翻译,一种非AUG机制)的二肽重复蛋白质(polyGA)的存在增加(数据未显示)。

[0986] 还生成F0小鼠,其小鼠C9orf72基因座被使用已知方法替换为含有GGGGCC(SEQ ID NO:1)的大约500次重复或600次重复的人类对应物。

[0987] 实施例4:靶向C9orf72-重复序列相邻内含子序列的siRNA降低含内含子的RNA及二肽重复蛋白质

[0988] 肌萎缩侧索硬化(ALS)及额颞叶退行(FTLD)是毁灭性神经退行性疾病,其在ALS情况下造成运动神经元疾病并且在FTLD情况下造成痴呆。两者都是致命性的。ALS及FTLD可作为自发性疾病或家族性(即,遗传性)疾病存在。ALS及FTLD的最常见遗传病因是C9orf72基因的5'非编码部分中六核苷酸重复序列(GGGGCC;SEQ ID NO:1)的扩增,该基因编码其功能尚未完全得以理解的蛋白质。未受影响的人通常已经在其C9orf72基因中具有数个及数十个六核苷酸重复序列,而那些发展出ALS及FTLD者从其父母的一遗传了数百至数千个六核苷酸重复序列拷贝的重复扩增。尚未有预测会造成C9orf72蛋白质的功能丧失的突变与ALS及FTLD关联。这些遗传性观察结果表明,C9orf72 ALS及FTLD是显性遗传疾病,导致病理学功能的增益。

[0989] C9orf72六核苷酸重复扩增如何造成运动神经元疾病及痴呆尚未可知,但C9orf72 ALS及FTLD患者的两种普遍尸检病理学发现与重复扩增相关联:(1)通过荧光原位杂交,可将含有正义及反义重复的RNA可视化为在神经元及其他细胞中截然不同的团簇;以及,(2)通过重复序列相关的非AUG依赖性翻译从该含正义及反义重复的RNA合成的二肽重复蛋白质,即,聚(甘氨酸-丙氨酸)、聚(谷氨酸-脯氨酸)、聚(甘氨酸-精氨酸)、聚(丙氨酸-脯氨酸)、及聚(脯氨酸-精氨酸),可通过免疫组织化学在细胞中检测。一种疾病假说提出,可视

化为团簇的含重复序列的RNA通过隔离RNA结合蛋白质而破坏细胞RNA代谢。第二种疾病假说认为，二肽重复蛋白对RNA代谢、蛋白质稳态及核细胞质转运具有广泛的毒性效应。两种病理学机制皆可能对疾病有所贡献。如果含有C9orf72重复序列的RNA转录本或其自身或作为二肽重复蛋白质翻译模板而促进ALS及FTLD中的致病性，则常规治疗策略将是摧毁含GGGGCC重复的RNA或根除其被翻译为二肽重复蛋白质的能力。

[0990] C9orf72基因从两个转录起始位点产生转录本。上游位点以候选的非编码外显子1A启动转录，而下游位点以候选的外显子1B启动转录。外显子1A及1B两者都可剪接至外显子2，外显子2含有蛋白质编码序列的起始位点。因为致病性六核苷酸重复扩增定位在外显子1A与外显子1B之间，从外显子1A起始至转录可产生含重复的RNA，而从外显子1B起始者则不产生。

[0991] 为了将C9orf72重复扩增疾病在小鼠中建模，将等位基因系列创建于小鼠胚胎干(ES)细胞中，其中，将来自人类C9orf72基因的片段(包括外显子1A的一部分、位于1A与1B之间的内含子、外显子1B的全部以及下游内含子的一部分)精确地放置在小鼠C9orf72基因中一个等位基因的其同源位置处。参见，例如，US 2018/0094267及WO 2018/064600，其各自通过援引而出于全部目的整体并入本文。将一系列GGGGCC重复扩增放置在于人类基因中发现的位置处，该位置从正常的三次重复到高达致病性的600次重复。参见实施例1。

[0992] 在培养中使携带不同重复扩增的小鼠ES细胞克隆分化为运动神经元，以研究该扩增对与ALS相关的细胞类型的效应。参见实施例3。在检查从经基因修饰的人源化C9orf72等位基因产生的转录本时发现，存在从三次重复的正常对照中占主导地位的外显子1B剪接转录本到具有更长重复扩增的等位基因中外显子1A剪接转录本的出现增加的转变。参见图12A、图12B、图12E及图12F。还观察到其丰度直接与六核苷酸重复长度相关的含未剪接内含子的转录本的累积，表明自我前馈环中的重复扩增越长，则自C9orf72基因产生的含重复序列的转录本越多。参见图12C、图12D、图12G及图12H。靶向含重复序列的内含子转录本以摧毁或激活二肽重复蛋白质合成的模板，同时保留正常C9orf72mRNA及蛋白质的合成，有望成为针对C9orf72重复扩增疾病的安全且有效的治疗策略。

[0993] 一种用以降低含C9orf72重复序列的RNA的可能途径是通过RNA干扰的天然过程实现，在该过程中，siRNA引导靶标RNA被RNA诱导沉默复合物直接裂解，之后RNA裂解片段被细胞核酸酶降解。然而，RNA干扰主要是细胞质过程，预期其不会作用于保留在细胞核中的RNA。含内含子的RNA通常是短寿命的，或作为mRNA前体，其被迅速剪接为成熟mRNA，或作为被剔除的内含子，于细胞核中迅速降解。因此，可合理地预期，含内含子的RNA将不能用于通过RNA干扰进行靶向。

[0994] 然而，在其中携带具有GGGGCC六核苷酸的300次重复的人源化C9orf72基因的鼠胚胎干细胞经靶向C9orf72 RNA序列的siRNA处理的实验中，已经证明，靶向邻近GGGGCC重复扩增的内含子序列的siRNA促使含内含子的C9orf72 RNA的累积降低至少60％(参见图16B)，而对于C9orf72成熟mRNA无效(参见图16A)。

[0995] 用于含内含子的RT-qPCR测定、测量含有邻近外显子1A的内含子序列的转录本的引物及探针，是来自表8的测定F。用于外显子2-3RT-qPCR测定的引物及探针购自ThermoFisher(测定法Mm01216829_m1)。对照GAPDH siRNA购自ThermoFisher(产品号4390849)。

[0996] 设计、合成了靶向C9orf72成熟mRNA或六核苷酸重复的C9orf72内含子5'的siRNA，并测定其在Neuro2a或BE(2)C细胞中的单剂量筛查的活性，如实施例5及6所描述。

[0997] 本实施例中，用来处理携带具有GGGGCC六核苷酸的300次重复的人源化C9orf72基因的小鼠胚胎干细胞的siRNA的正义链及反义链的未经修饰的核苷酸序列显示于表10A中。本实施例中，用来处理携带具有GGGGCC六核苷酸序列的300次重复的人源化C9orf72基因的小鼠胚胎干细胞的siRNA的正义链及反义链的经修饰的核苷酸序列显示于表10B中。

[0998] 靶向内含子的siRNA也降低二肽重复蛋白质的产生，降低幅度多达75％。参见图17A至图17B。这些出乎意料的实验结果指示，细胞中累积的具有C9orf72六核苷酸重复扩增的含内含子的RNA易受RNA干扰的影响。结果表明，负责二肽重复蛋白质合成的相当一部分含内含子的C9orf72 RNA位于细胞质中。相比之下，靶向C9orf72 mRNA蛋白质编码序列的siRNA产生mRNA的强烈敲低，但对于含内含子的转录本无效，并且未明显降低二肽重复蛋白质合成。参见图16A至图16B。靶向内含子的siRNA与靶向mRNA的siRNA之间的结果发散性表明，两类靶向序列存在于未共价连接的独立RNA上上。

[0999] 表10A.C9orf72 siRNA.

[1000]

#	双链体 ID	靶向区域	正义链序列 5' 至 3'	反义链序列 5' 至 3'
1	AD-348904	外显子 11	GAAGACCUUUCUACACUAGUU (SEQ ID NO: 14)	AACUAGTGUAGAAAGGUCUUCCA (SEQ ID NO: 15)
2	AD-348136	外显子 5	UAGCUGAUACAGUACUCAAUU (SEQ ID NO: 16)	AAUUGAGUACUGUAUCAGCUAUA (SEQ ID NO: 17)
3	AD-347612	外显子 2	CAAGACAGAGAUUGCUUUAAU (SEQ ID NO: 18)	AUUAAAGCAAUCUCUGUCUUGGC (SEQ ID NO: 19)
4	AD-348639	外显子 9	GUCUUACACAGAGACACUCUA (SEQ ID NO: 20)	UAGAGUGUCUCUGUGUAAGACAU (SEQ ID NO: 21)
5	AD-347606	外显子 2	UGUUGCCAAGACAGAGAUUGU (SEQ ID NO: 22)	ACAAUCTCUGUCUUGGCAACAGC (SEQ ID NO: 23)
6	AD-348140	外显子 5	UGAUACAGUACUCAAUGAUGA (SEQ ID NO: 24)	UCAUCATUGAGUACUGUAUCAGC (SEQ ID NO: 25)
7	AD-463863	内含子 1	AAAGACCUGAUAAAGAUUAAU (SEQ ID NO: 2)	AUUAAUCUUUAUCAGGUCUUUUC (SEQ ID NO: 3)
8	AD-463862	内含子 1	AAAAGACCUGAUAAAGAUUAA (SEQ ID NO: 4)	UUAAUCUUUAUCAGGUCUUUUCU (SEQ ID NO: 5)
9	AD-463869	内含子 1	CUGAUAAAGAUUAACCAGAAU (SEQ ID NO: 6)	AUUCUGGUUAAUCUUUAUCAGGU (SEQ ID NO: 7)
10	AD-463873	内含子 1	AAAGAUUAACCAGAAGAAAU (SEQ ID NO: 8)	AUUUUCUUCUGGUUAAUCUUUAU (SEQ ID NO: 9)
11	AD-463872	内含子 1	AUAAAGAUUAACCAGAAGAAA (SEQ ID NO: 10)	UUUCUUCUGGUUAAUCUUUAUCA (SEQ ID NO: 11)
12	AD-463860	内含子 1	AGAAAAGACCUGAUAAAGAUU (SEQ ID NO: 12)	AAUCUUUAUCAGGUCUUUUCUUG (SEQ ID NO: 13)

[1001] 表10B.C9orf72 siRNA.

[1002]

双链体 ID	名称	靶标	链	序列 5’至 3’
AD-347606	siRNA #5	外显子 2	正义 反义	usgsuugcCfaAfGfAfcagagauuguL96 asCfsaauc(Tgn)cugucuUfgGfcaacasgsc
AD-347612	siRNA #3	外显子 2	正义 反义	csasagacAfgAfGfAfuugcuuuaauL96 asUfsuaaa(Ggn)caaucuCfuGfucuugsgsc
AD-348136	siRNA #2	外显子 5	正义 反义	usasgcugAfuAfCfAfguacucaauuL96 asAfsuuga(Ggn)uacuguAfuCfagcuasusa
AD-348140	siRNA #6	外显子 5	正义 反义	usgsauacAfgUfAfCfucaaugaugaL96 usCfsauca(Tgn)ugaguaCfuGfuaucasgsc
AD-348639	siRNA #4	外显子 9	正义 反义	gsuscuuaCfaCfAfGfagacacucuaL96 usAfsgagu(Ggn)ucucugUfgUfaagacsasu
AD-348904	siRNA #1	外显子 11	正义 反义	gsasagacCfuUfUfCfuacacuaguuL96 asAfscuag(Tgn)guagaaAfgGfucuucscsa
AD-463860	siRNA #12	内含子 1A	正义 反义	asgsaaaaGfaCfCfUfgauaaagauuL96 asAfsucuUfuAfUfcaggUfcUfuuucususg
AD-463862	siRNA #8	内含子 1A	正义 反义	asasaagaCfcUfGfAfuaaagauuaaL96 usUfsaauCfuUfUfaucaGfgUfcuuuuscsu
AD-463863	siRNA #7	内含子 1A	正义 反义	asasagacCfuGfAfUfaaagauuaauL96 asUfsuaaUfcUfUfuaucAfgGfucuuususc
AD-463869	siRNA #9	内含子 1A	正义 反义	csusgauaAfaGfAfUfuaaccagaauL96 asUfsucuGfgUfUfaaucUfuUfaucagsgsu
AD-463872	siRNA #11	内含子 1A	正义 反义	asusaaagAfuUfAfAfccagaagaaaL96 usUfsucuUfcUfGfguuaAfuCfuuuauscsa
AD-463873	siRNA #10	内含子 1A	正义 反义	asasagauUfaAfCfCfagaagaaaauL96 asUfsuuuCfuUfCfugguUfaAfucuuusasu

[1003] SEQ ID NO:2、4、6、8、10、12、14、16、18、20、22及24的正义链包含全部下列特征/修饰：(a)21个核苷酸的长度；(b)位于位置1至6、8及12至21(自5’端计数)处的2’-O-甲基修饰；(c)位于位置7及9至11(自5’端计数)处的2’-F修饰；(d)位于位置1与2之间及核苷酸位置2与3之间(自5’端计数)的硫代磷酸酯核苷酸间键；以及(e)附接至3’端的去唾液酸糖蛋白受体(ASGPR)配体，其中，ASGPR配体包含通过三价分支接头附接的GalNAc衍生物。

[1004] SEQ ID NO:3、5、7、9、11及13的反义链包含全部下列特征/修饰：(a)23个核苷酸的长度；(b)位于核苷酸位置1、3至5、7、10至13、15及17至23(自5’端计数)处的2’-O-甲基修饰；(c)位于核苷酸位置2、6、8、9、14及16(自5’端计数)处的2’-F修饰；以及(d)位于核苷酸位置1与2之间、核苷酸位置2与3之间、核苷酸位置21与22之间及核苷酸位置22与23之间(自5’端计数)的硫代磷酸酯核苷酸间键。

[1005] SEQ ID NO:15、17、19、21、23及25的反义链包含全部下列特征/修饰：(a)23个核苷酸的长度；(b)位于位置1、3至6、8至13、15及17至23处的2’-O-甲基修饰；(c)位于位置2、14及16处的2’-F修饰；(d)位于位置7处的二醇核酸(GNA)；以及，(e)位于核苷酸位置1与2之间、核苷酸位置2与3之间、核苷酸位置21与22之间及核苷酸位置22与23之间(自5’端计数)的硫代磷酸酯核苷酸间键。

[1006] 表10A及10B中的双链RNAi剂各自包含位于该反义链的3’末端的具有2个核苷酸的突出，以及位于该反义链的5’末端的钝端。

[1007] 实施例5:RNAi剂设计、合成、选择及体外评估

[1008] 本实施例描述C9orf72 RNAi剂的设计、合成、选择及体外评估的方法。

[1009] 试剂的来源

[1010] 若本文中未具体给出试剂的来源，则测量试剂可从任何分子生物学用试剂供货商以用于分子生物学应用的质量/纯度标准获得。

[1011] 生物信息学

[1012] 使用定制的R及Python脚本设计靶向人类C9orf72转录本(C9orf72；人NCBI refseqID NM_001256054;NCBI GeneID:203228)的siRNA。人NM_001256054mRNA的长度为3356个碱基。

[1013] 未修饰的C9orf72正义链及反义链核苷酸序列的详细列述显示于表12中。经修饰的C9orf72正义链及反义链核苷酸序列的详细列述显示于表13中。

[1014] 使用定制R及Python脚本，设计了靶向人类C9orf72基因(GenBank登录号NC_000009.12)的位于外显子1a与1b之间的内含子的siRNA。未修饰的C9orf72正义链及反义链核苷酸序列的详细列述显示于表14及19中。经修饰的C9orf72正义链及反义链核苷酸序列的详细列述显示于表15及20中。

[1015] 应理解，本申请通篇中，不具小数的双链体名称等同于具有小数的双链体名称，仅引用双链体的批号。例如，AD-347430等同于AD-347430.1。

[1016] 体外Cos-7(双萤光素酶psiCHECK2载体)、BE(2)-C及Neuro-2a筛查

[1017] 细胞培养及转染：

[1018] 通过如下转染Cos-7(ATCC)：于384孔板中，将每孔5μl的稀释于Opti-MEM中的1ng/ul的C9orf72内含子1psiCHECK2载体(Blue Heron Biotechnology)、4.9μl的Opti-MEM加上0.1μl的Lipofectamine 2000(Invitrogen,Carlsbad CA.，目录号11668-019)添加到每孔5μl的siRNA双链体中，且每一siRNA双链体存在4组重复，以及在室温下温育15分钟。随后，将含有～5x10^3个细胞的35μl的杜尔贝科改良伊格尔培养基(ThermoFisher)加至该siRNA混合物中。将细胞温育48小时，之后进行萤火虫萤光素酶(转染对照)及海肾萤光素酶(融合至靶标序列)测量。执行三个剂量的实验，分别为10nM、1nM及0.1nM。

[1019] 通过如下转染BE(2)-C(ATCC)：于384孔板中，将每孔4.9μl的Opti-MEM加上0.1μl的RNAiMAX(Invitrogen,Carlsbad CA.，目录号13778-150)添加至每孔5μl的siRNA双链体中，且每一siRNA双链体存在4组重复，以及在室温下温育15分钟。随后，将含有～5x10^3个细胞的40μl的最小必需培养基与F12培养基的1:1混合物(ThermoFisher)加至siRNA混合物中。在RNA纯化之前，将细胞温育48小时。执行两个剂量的实验，分别为10nM及0.1nM。

[1020] 通过如下转染Neuro-2a(ATCC)：于384孔板中，将每孔4.9μl的Opti-MEM加上0.1μl的RNAiMAX(Invitrogen,Carlsbad CA.，目录号13778-150)添加至每孔5μl的siRNA双链体中，且每一siRNA双链体存在4组重复，以及在室温下温育15分钟。随后，将含有～5x10^3个细胞的40μl的最小必需培养基(ThermoFisher)加至siRNA混合物中。在RNA纯化之前，将细胞温育48小时。执行两个剂量的实验，分别为10nM及0.1nM。

[1021] 使用DYNABEADS mRNA分离试剂盒进行的总RNA分离：

[1022] 使用自动方法在BioTek-EL406平台上使用DYNABEADs(Invitrogen,cat#61012)分离RNA。简而言之，将70ul的裂解液/结合缓冲液及10ul的含有33μl磁性微珠的裂解缓冲液加至具有细胞的板中。将板于电磁培养箱中于室温温育10分钟，随后捕获磁性微珠并移除

上清液。随后，将微珠结合的RNA以150μl洗涤缓冲液A洗涤2次，以洗涤缓冲液B洗涤一次。将微珠以150μl洗脱缓冲液洗涤，再次捕获并移除上清液。

[1023] 使用ABI高能cDNA逆转录试剂盒(Applied Biosystems,Foster City,CA,Cat#4368813)的cDNA合成：

[1024] 将10μl的含有1μl 10X缓冲液、0.4ul 25X dNTPs、1μl 10x随机引物、0.5μl逆转录酶、0.5μl Rnase抑制剂及6.6μl H2O每反应的预混液加至上述分离的RNA中。将板密封、混合、并在电磁培养箱中于室温温育10分钟，之后于37℃温育2h。

[1025] 实时PCR：

[1026] 在384孔板(Roche，目录号04887301001)的各孔内，将2μl的cDNA及5μl的Lightcycler 480探针预混物(Roche，目录号04887301001)加至0.5μl的人GAPDH TaqMan探针(4326317E)及0.5μl的C9orf72人类探针(Hs00376619_m1,Thermo)中，或加至0.5μl的小鼠GAPDH TaqMan探针(4352339E)及0.5μl的C9orf72小鼠探针(Mm01216837_m1,Thermo)中。实时PCR是于LightCycler480实时PCR系统(Roche)中进行。每一双链体测试至少两次，且数据相对于使用非靶向对照siRNA转染的细胞归一化。为了计算相对倍数改变，使用ΔΔCt方法分析数据并将该数据相对于使用以非靶向对照siRNA转染的细胞实施的测定归一化。

[1027] 在Neuro2a细胞中进行的表12及表13所列dsRNA剂的筛查结果显示于表16中。在BE(2)c细胞中进行的表12及表13所列dsRNA剂的筛查结果显示于表17中。在Cos7细胞中进行的表14及表15所列dsRNA剂的筛查结果显示于表18中。

[1028] 表11.核苷酸序列呈现中使用的核苷酸单体的缩写。应理解，当这些单体存在于寡核苷酸中时，它们借由5'-3'-磷酸二酯键而相互连接。

[1029]

缩写	核苷酸
A	腺苷-3'-磷酸酯
Ab	β-L-腺苷-3`-磷酸酯
Abs	β-L-腺苷-3`-硫代磷酸酯
Af	2'-氟腺苷-3'-磷酸酯
Afs	2'-氟腺苷-3'-硫代磷酸酯
As	腺苷-3'-硫代磷酸酯
C	胞苷-3'-磷酸酯
Cb	β-L-胞苷-3`-磷酸酯
Cbs	β-L-胞苷-3'-硫代磷酸酯
Cf	2'-氟胞苷-3'-磷酸酯

[1030]

缩写	核苷酸
Cfs	2'-氟胞苷-3'-硫代磷酸酯
Cs	胞苷-3'-硫代磷酸酯
G	鸟苷-3'-磷酸酯
Gb	β-L-鸟苷-3`-磷酸酯
Gbs	β-L-鸟苷-3`-硫代磷酸酯
Gf	2'-氟鸟苷-3'-磷酸酯
Gfs	2'-氟鸟苷-3'-硫代磷酸酯
Gs	鸟苷-3'-硫代磷酸酯
T	5'-甲基尿苷-3'-磷酸酯
Tf	2'-氟-5-甲基尿苷-3'-磷酸酯
Tfs	2'-氟-5-甲基尿苷-3'-硫代磷酸酯
Ts	5-甲基尿苷-3'-硫代磷酸酯
U	尿苷-3'-磷酸酯
Uf	2'-氟尿苷-3'-磷酸酯
Ufs	2'-氟尿苷 -3'-硫代磷酸酯
Us	尿苷 -3'-硫代磷酸酯
N	任意核苷酸，经修饰或未修饰
a	2'-O-甲基腺苷-3'-磷酸酯
as	2'-O-甲基腺苷-3'- 硫代磷酸酯
c	2'-O-甲基胞苷-3'-磷酸酯
cs	2'-O-甲基胞苷-3'- 硫代磷酸酯
g	2'-O-甲基鸟苷-3'-磷酸酯
gs	2'-O-甲基鸟苷-3'- 硫代磷酸酯
t	2'-O-甲基-5-甲基尿苷-3'-磷酸酯
ts	2'-O-甲基-5-甲基尿苷-3'-硫代磷酸酯
u	2'-O-甲基尿苷-3'-磷酸酯
us	2'-O-甲基尿苷-3'-硫代磷酸酯
s	硫代磷酸酯键
L96	N-[三(GalNAc-烷基)-酰胺基癸酰基)]-4-羟基脯氨醇 Hyp-(GalNAc-烷基)3
Y34	2-羟基甲基-四氢呋喃-4-甲氧基-3-磷酸酯(无碱基的 2'-OMe 呋喃糖)
Y44	反向的无碱基 DNA (2-羟基甲基-四氢呋喃-5-磷酸酯)
(Agn)	腺苷-二醇核酸(GNA)
(Cgn)	胞苷-二醇核酸(GNA)
(Ggn)	鸟苷-二醇核酸(GNA)
(Tgn)	胸苷-二醇核酸(GNA) S-异构体
P	磷酸酯
VP	乙烯基-膦酸酯
dA	2'-脱氧腺苷-3'-磷酸酯
dAs	2'-脱氧腺苷-3'-硫代磷酸酯
dC	2'-脱氧胞苷-3'-磷酸酯
dCs	2'-脱氧胞苷-3'-硫代磷酸酯

缩写	核苷酸
dG	2'-脱氧鸟苷-3'-磷酸酯
dGs	2'-脱氧鸟苷-3'-硫代磷酸酯
dT	2'-脱氧胸苷-3'-磷酸酯
dTs	2'-脱氧胸苷-3'-硫代磷酸酯
dU	2'-脱氧尿苷
dUs	2`-脱氧尿苷-3`-硫代磷酸酯
(Ahd)	2'-O-十六烷基-腺苷-3'-磷酸酯
(Ahds)	2'-O-十六烷基-腺苷-3'-硫代磷酸酯
(Chd)	2'-O-十六烷基-胞苷-3'-磷酸酯
(Chds)	2'-O-十六烷基-胞苷-3'-硫代磷酸酯
(Ghd)	2'-O-十六烷基-鸟苷-3'-磷酸酯
(Ghds)	2'-O-十六烷基-鸟苷-3'-硫代磷酸酯
(Uhd)	2'-O-十六烷基-尿苷-3'-磷酸酯
(Uhds)	2'-O-十六烷基-尿苷-3'-硫代磷酸酯

[1031]

[1032]

表 12. 靶向 C9orf72 编码区的 dsRNA 剂的未修饰正义及反义链序列

双链体名称	正义序列 5'至 3'	SEQ ID NO:	来源及范围	NM_001256054.2 中范围	翻译序列 5'至 3'	SEQ ID NO:	来源及范围	NM_001256054.2 中范围
AD-347430.1	UGCGUCAAACAGCGA CAAGUU		NM_001256054.2_3-55_s	35-55	AACUUGUCGCUGUUUGAC GCACC		NM_001256054.2_33-55_as	33-55
AD-347432.1	CGUCAAACAGCGACA AGUUCU		NM_001256054.2_37-57_C21U_s	37-57	AGAACUUGUCGCUGUUUG ACGCA		NM_001256054.2_35-57_G1A_as	35-57
AD-347454.1	CCCACGUAAAAGAUG ACGCUU		NM_001256054.2_59-79_s	59-79	AAGCGUCAUCUUUUACGU GGGCG		NM_001256054.2_57-79_as	57-79
AD-347455.1	CCACGUAAAAGAUGA CGCUUU		NM_001256054.2_60-80_G21U_s	60-80	AAAGCGUCAUCUUUUACG UGGGC		NM_001256054.2_58-80_C1A_as	58-80
AD-347570.1	AGUGAUGUCGACUCU UUGCCU		NM_001256054.2_199-219_C21U_s	199-219	AGGCAAAGAGUCGACAUC ACUGC		NM_001256054.2_197-219_G1A_as	197-219
AD-347602.1	CAGCUGUUGCCAAGA CAGAGA		NM_001256054.2_231-251_s	231-251	UCUCUGTCUUGGCAACAG CUGGA		NM_001256054.2_229-251_as	229-251
AD-347603.1	AGCUGUUGCCAAGAC AGAGAU		NM_001256054.2_232-252_s	232-252	AUCUCUGUCUUGGCAACA GCUGG		NM_001256054.2_230-252_as	230-252
AD-347606.1	UGUUGCCAAGACAGA GAUUGU		NM_001256054.2_235-255_C21U_s	235-255	ACAAUCTCUGUCUUGGCA ACAGC		NM_001256054.2_233-255_G1A_as	233-255
AD-347610.1	GCCAAGACAGAGAUU GCUUUA		NM_001256054.2_239-259_s	239-259	UAAAGCAAUCUCUGUCUU GGCAA		NM_001256054.2_237-259_as	237-259
AD-347612.1	CAAGACAGAGAUUGC UUUAAU		NM_001256054.2_241-261_G21U_s	241-261	AUUAAAGCAAUCUCUGUC UUGGC		NM_001256054.2_239-261_C1A_as	239-261

[1033]

AD-347613.1	AAGACAGAGAUUGCU UUAAGU		NM_001256054.2_2 42-262_s	242-262	ACUUAAAGCAAUCUCUGU CUUGG		NM_001256054.2_240-262_as	240-262
AD-347614.1	AGACAGAGAUUGCUU UAAGUU		NM_001256054.2_2 43-263_G21U_s	243-263	AACUUAAAGCAAUCUCUG UCUUG		NM_001256054.2_241-263_C1A_as	241-263
AD-347615.1	GACAGAGAUUGCUUU AAGUGU		NM_001256054.2_2 44-264_G21U_s	244-264	ACACUUAAAGCAAUCUCU GUCUU		NM_001256054.2_242-264_C1A_as	242-264
AD-347650.1	UAGCAGCUACUUUUG CUUACU		NM_001256054.2_2 79-299_s	279-299	AGUAAGCAAAAGUAGCUG CUAAU		NM_001256054.2_277-299_as	277-299
AD-347674.1	ACAAUAUUCUUGGUC CUAGAU		NM_001256054.2_3 03-323_G21U_s	303-323	AUCUAGGACCAAGAAUAU UGUCC		NM_001256054.2_301-323_C1A_as	301-323
AD-347676.1	AAUAUUCUUGGUCCU AGAGUA		NM_001256054.2_3 05-325_s	305-325	UACUCUAGGACCAAGAAU AUUGU		NM_001256054.2_303-325_as	303-325
AD-347677.1	AUAUUCUUGGUCCUA GAGUAA		NM_001256054.2_3 06-326_s	306-326	UUACUCTAGGACCAAGAA UAUUG		NM_001256054.2_304-326_as	304-326
AD-347716.1	AGACAGAACAGGUAC UUCUCA		NM_001256054.2_3 45-365_s	345-365	UGAGAAGUACCUGUUCUG UCUUU		NM_001256054.2_343-365_as	343-365
AD-347773.1	UAAAUGGAGAAAUCC UUCGAA		NM_001256054.2_4 02-422_s	402-422	UUCGAAGGAUUUCUCCAU UUAGA		NM_001256054.2_400-422_as	400-422
AD-347778.1	GGAGAAAUCCUUCGA AAUGCA		NM_001256054.2_4 07-427_s	407-427	UGCAUUTCGAAGGAUUUC UCCAU		NM_001256054.2_405-427_as	405-427
AD-347842.1	CAUUAAUCUUUGAUG GAAACU		NM_001256054.2_4 89-509_s	489-509	AGUUUCCAUCAAAGAUUA AUGAA		NM_001256054.2_487-509_as	487-509
AD-347843.1	AUUAAUCUUUGAUG GAAACUU		NM_001256054.2_4 90-510_G21U_s	490-510	AAGUUUCCAUCAAAGAUU AAUGA		NM_001256054.2_488-510_C1A_as	488-510
AD-347863.1	CAUAUGGACUAUCAA UUAUAU		NM_001256054.2_5 28-548_C21U_s	528-548	AUAUAATUGAUAGUCCAU AUGUG		NM_001256054.2_526-548_G1A_as	526-548

[1034]

AD-347865.1	UAUGGACUAUCAAUU AUACUU		NM_001256054.2_5 30-550_s	530-550	AAGUAUAAUUGAUAGUCC AUAUG		NM_001256054.2_528-550_as	528-550
AD-347867.1	UGGACUAUCAAUAU ACUUCU		NM_001256054.2_5 32-552_C21U_s	532-552	AGAAGUAUAAUUGAUAGU CCAUA		NM_001256054.2_530-552_G1A_as	530-552
AD-347874.1	UCAAUUAUACUUCCA CAGACA		NM_001256054.2_5 39-559_s	539-559	UGUCUGTGGAAGUAUAAU UGAUA		NM_001256054.2_537-559_as	537-559
AD-347892.1	ACAGAACUUAGUUUC UACCUU		NM_001256054.2_5 57-577_C21U_s	557-577	AAGGUAGAAACUAAGUUC UGUCU		NM_001256054.2_555-577_G1A_as	555-577
AD-347893.1	CAGAACUUAGUUUCU ACCUCU		NM_001256054.2_5 58-578_C21U_s	558-578	AGAGGUAGAAACUAAGUU CUGUC		NM_001256054.2_556-578_G1A_as	556-578
AD-347923.1	GAGUGUGUUGAU AGAUUAA		NM_001256054.2_5 88-608_s	588-608	UUAAUCTAUCAACACACAC UCUA		NM_001256054.2_586-608_as	586-608
AD-347924.1	AGUGUGUUGAUA GAUUAAU		NM_001256054.2_5 89-609_C21U_s	589-609	AUUAAUCUAUCAACACAC ACUCU		NM_001256054.2_587-609_G1A_as	587-609
AD-347926.1	UGUGUGUUGAUAGA UUAACAU		NM_001256054.2_5 91-611_C21U_s	591-611	AUGUUAAUCUAUCAACAC ACACU		NM_001256054.2_589-611_G1A_as	589-611
AD-347942.1	AACACAUAUAAUCCG GAAAGU		NM_001256054.2_6 07-627_G21U_s	607-627	ACUUUCCGGAUUAUAUGU GUUAA		NM_001256054.2_605-627_C1A_as	605-627
AD-347993.1	UGUCCAGAAGAUUAU CUUAGA		NM_001256054.2_6 61-681_s	661-681	UCUAAGAUAAUCUUCUGG ACAUU		NM_001256054.2_659-681_as	659-681
AD-348087.1	GAACUGCUUUCAUCU AUGAAA		NM_001256054.2_7 55-775_s	755-775	UUUCAUAGAUGAAAGCAG UUCCA		NM_001256054.2_753-775_as	753-775
AD-348136.1	UAGCUGAUACAGUAC UCAAUU		NM_001256054.2_8 04-824_G21U_s	804-824	AAUUGAGUACUGUAUCAG CUAUA		NM_001256054.2_802-824_C1A_as	802-824
AD-348140.1	UGAUACAGUACUCAA UGAUGA		NM_001256054.2_8 08-828_s	808-828	UCAUCATUGAGUACUGUA UCAGC		NM_001256054.2_806-828_as	806-828

[1035]

AD-348176.1	CUGUCAUGAAGGCUU UCUUCU		NM_001256054.2_844-864_s	844-864	AGAAGAAAGCCUUCAUGA CAGCU		NM_001256054.2_842-864_as	842-864
AD-348231.1	UGUUCCGUUGUAGU AGGUAGU		NM_001256054.2_899-919_C21U_s	899-919	ACUACCTACUACAACGGAA CAGC		NM_001256054.2_897-919_G1A_as	897-919
AD-348341.1	AUCAUUUAAAUAUGA GUCAGU		NM_001256054.2_1012-1032_G21U_s	1012-1032	ACUGACTCAUAUUUAAAU GAUGA		NM_001256054.2_1010-1032_C1A_as	1010-1032
AD-348377.1	CCUGCUAAAGGAUUC AACUGU		NM_001256054.2_1048-1068_G21U_s	1048-1068	ACAGUUGAAUCCUUUAGC AGGCC		NM_001256054.2_1046-1068_C1A_as	1046-1068
AD-348487.1	CACCCUGUCAUGAAC AUAUUU		NM_001256054.2_1158-1178_s	1158-1178	AAAUAUGUUCAUGACAGG GUGGC		NM_001256054.2_1156-1178_as	1156-1178
AD-348497.1	UGAACAUAUUUUAUAA UCAGCU		NM_001256054.2_1168-1188_G21U_s	1168-1188	AGCUGATUAUAAAUAUGU UCAUG		NM_001256054.2_1166-1188_C1A_as	1166-1188
AD-348500.1	ACAUAUUUAUAAUCA GCGUAU		NM_001256054.2_1171-1191_G21U_s	1171-1191	AUACGCTGAUUAUAAAUA UGUUC		NM_001256054.2_1169-1191_C1A_as	1169-1191
AD-348502.1	AUAUUUAUAAUCAGC GUAGAU		NM_001256054.2_1173-1193_s	1173-1193	AUCUACGCUGAUUAUAAA UAUGU		NM_001256054.2_1171-1193_as	1171-1193
AD-348578.1	UCAGGAUACGAUCAU CUACAU		NM_001256054.2_1249-1269_C21U_s	1249-1269	AUGUAGAUGAUCGUAUCC UGAGC		NM_001256054.2_1247-1269_G1A_as	1247-1269
AD-348588.1	AUCAUCUACACUGAC GAAAGU		NM_001256054.2_1259-1279_C21U_s	1259-1279	ACUUUCGUCAGUGUAGAU GAUCG		NM_001256054.2_1257-1279_G1A_as	1257-1279
AD-348590.1	CAUCUACACUGACGA AAGCUU		NM_001256054.2_1261-1281_s	1261-1281	AAGCUUTCGUCAGUGUAG AUGAU		NM_001256054.2_1259-1281_as	1259-1281
AD-348591.1	AUCUACACUGACGAA AGCUUU		NM_001256054.2_1262-1282_s	1262-1282	AAAGCUTUCGUCAGUGUA GAUGA		NM_001256054.2_1260-1282_as	1260-1282
AD-348596.1	CACUGACGAAAGCUU UACUCU		NM_001256054.2_1267-1287_C21U_s	1267-1287	AGAGUAAAGCUUUCGUCA GUGUA		NM_001256054.2_1265-1287_G1A_as	1265-1287

[1036]

AD-348597.1	ACUGACGAAAAGCUUU ACUCCU	NM_001256054.2_1 268-1288_s	1268-1288	AGGAGUAAAGCUUUCGUC AGUGU		NM_001256054.2_1266 -1288_as	1266-1288
AD-348598.1	CUGACGAAAAGCUUUA CUCCUU	NM_001256054.2_1 269-1289_G21U_s	1269-1289	AAGGAGTAAAGCUUUCGU CAGUG		NM_001256054.2_1267 -1289_C1A_as	1267-1289
AD-348630.1	UUUCAAGAUGUCUUA CACAGA	NM_001256054.2_1 301-1321_s	1301-1321	UCUGUGTAAGACAUCUUG AAAAA		NM_001256054.2_1299 -1321_as	1299-1321
AD-348639.1	GUCUUACACAGAGAC ACUCUA	NM_001256054.2_1 310-1330_s	1310-1330	UAGAGUGUCUCUGUGUA AGACAU		NM_001256054.2_1308 -1330_as	1308-1330
AD-348672.1	CUGGAUCAGGUCUUU CAGCUU	NM_001256054.2_1 343-1363_G21U_s	1343-1363	AAGCUGAAAGACCUGAUC CAGGA		NM_001256054.2_1341 -1363_C1A_as	1341-1363
AD-348708.1	CUCAGAAGUACUUUC CUUGCA	NM_001256054.2_1 379-1399_s	1379-1399	UGCAAGGAAAGUACUUCU GAGAG		NM_001256054.2_1377 -1399_as	1377-1399
AD-348734.1	UCUACUUGUCCUUCA CAGAAA	NM_001256054.2_1 405-1425_s	1405-1425	UUUCUGTGAAGGACAAGU AGAAA		NM_001256054.2_1403 -1425_as	1403-1425
AD-348751.1	GAAAAGCCUUGACAC UAAUAA	NM_001256054.2_1 422-1442_s	1422-1442	UUAUUAGUGUCAAGGCU UUUCUG		NM_001256054.2_1420 -1442_as	1420-1442
AD-348787.1	CCUUUAAAUCUCUUC GGAACU	NM_001256054.2_1 479-1499_C21U_s	1479-1499	AGUUCCGAAGAGAUUUAA AGGGC		NM_001256054.2_1477 -1499_G1A_as	1477-1499
AD-348788.1	CUUUAAAUCUCUUCG GAACCU	NM_001256054.2_1 480-1500_s	1480-1500	AGGUUCCGAAGAGAUUUA AAGGG		NM_001256054.2_1478 -1500_as	1478-1500
AD-348789.1	UUUAAAUCUCUUCGG AACCUU	NM_001256054.2_1 481-1501_G21U_s	1481-1501	AAGGUUCCGAAGAGAUUU AAAGG		NM_001256054.2_1479 -1501_C1A_as	1479-1501
AD-348836.1	GGGCGAUCUUAACAU AAUAAU	NM_001256054.2_1 528-1548_s	1528-1548	AUUAUUAUGUUAAGAUCG CCCUC		NM_001256054.2_1526 -1548_as	1526-1548
AD-348842.1	UCUUAACAUAAUAAU GGCUCU	NM_001256054.2_1 534-1554_s	1534-1554	AGAGCCAUUAUUAUGUUA AGAUC		NM_001256054.2_1532 -1554_as	1532-1554

[1037]

AD-348895.1	UUAUCUUUGGAAGAC CUUUCU		NM_001256054.2_1 587-1607_s	1587-1607	AGAAAGGUCUUCCAAAGA UAAAA		NM_001256054.2_1585 -1607_as	1585-1607
AD-348904.1	GAAGACCUUUCUACA CUAGUU		NM_001256054.2_1 596-1616_G21U_s	1596-1616	AACUAGTGUAGAAAGGUC UUCCA		NM_001256054.2_1594 -1616_C1A_as	1594-1616
AD-348905.1	AAGACCUUUCUACAC UAGUGU		NM_001256054.2_1 597-1617_s	1597-1617	ACACUAGUGUAGAAAGGU CUUCC		NM_001256054.2_1595 -1617_as	1595-1617
AD-348906.1	AGACCUUUCUACACU AGUGUU		NM_001256054.2_1 598-1618_G21U_s	1598-1618	AACACUAGUGUAGAAAGG UCUUC		NM_001256054.2_1596 -1618_C1A_as	1596-1618
AD-348926.1	GCAAGAACGAGAUGU UCUAAU		NM_001256054.2_1 618-1638_s	1618-1638	AUUAGAACACUCUCGUUCU UGCAC		NM_001256054.2_1616 -1638_as	1616-1638
AD-348930.1	GAACGAGAUGUUCUA AUGACU		NM_001256054.2_1 622-1642_s	1622-1642	AGUCAUTAGAACACUCUCG UUCUU		NM_001256054.2_1620 -1642_as	1620-1642
AD-348931.1	AACGAGAUGUUCUAA UGACUU		NM_001256054.2_1 623-1643_s	1623-1643	AAGUCATUAGAACACUCUC GUUCU		NM_001256054.2_1621 -1643_as	1621-1643
AD-348932.1	ACGAGAUGUUCUAAU GACUUU		NM_001256054.2_1 624-1644_s	1624-1644	AAAGUCAUUAGAACACUCU CGUUC		NM_001256054.2_1622 -1644_as	1622-1644
AD-348957.1	AUGUGUAACUUAAUA AGCCUA		NM_001256054.2_1 649-1669_s	1649-1669	UAGGCUTAUUAAGUUACA CAUUU		NM_001256054.2_1647 -1669_as	1647-1669
AD-348959.1	GUGUAACUUAAUAAG CCUAUU		NM_001256054.2_1 651-1671_s	1651-1671	AAUAGGCUUAUUAAGUUA CACAU		NM_001256054.2_1649 -1671_as	1649-1671
AD-348961.1	GUAACUUAAUAAGCC UAUUCU		NM_001256054.2_1 653-1673_C21U_s	1653-1673	AGAAUAGGCUUAUUAAGU UACAC		NM_001256054.2_1651 -1673_G1A_as	1651-1673
AD-348962.1	UAACUUAAUAAGCCU AUUCCA		NM_001256054.2_1 654-1674_s	1654-1674	UGGAAUAGGCUUAUUAAG UUACA		NM_001256054.2_1652 -1674_as	1652-1674
AD-348963.1	AACUUAAUAAGCCUA UUCCAU		NM_001256054.2_1 655-1675_s	1655-1675	AUGGAATAGGCUUAUUAA GUUAC		NM_001256054.2_1653 -1675_as	1653-1675

[1038]

AD-349042.1	GUUAAGUAAGUUACA CUACAU	NM_001256054.2_1 765-1785_G21U_s	1765-1785	AUGUAGTGUAACUUACUU AACUG		NM_001256054.2_1763 -1785_C1A_as	1763-1785
AD-349048.1	UAAGUUACACUACAG UUCUCA	NM_001256054.2_1 771-1791_s	1771-1791	UGAGAACUGUAGUGUAAC UUACU		NM_001256054.2_1769 -1791_as	1769-1791
AD-349051.1	GUUACACUACAGUUC UCACAA	NM_001256054.2_1 774-1794_s	1774-1794	UUGUGAGAACUGUAGUG UAACUU		NM_001256054.2_1772 -1794_as	1772-1794
AD-349126.1	CAGACCUAUGUUUAC AAUAUA	NM_001256054.2_1 867-1887_s	1867-1887	UAUAUUGUAAACAUAGGU CUGUA		NM_001256054.2_1865 -1887_as	1865-1887
AD-349127.1	AGACCUAUGUUUACA AUAUAA	NM_001256054.2_1 868-1888_s	1868-1888	UUAUAUTGUAAACAUAGG UCUGU		NM_001256054.2_1866 -1888_as	1866-1888
AD-349395.1	GAUCAAGCAGAUGUU UAAUUU	NM_001256054.2_2 165-2185_G21U_s	2165-2185	AAAUUAAACAUCUGCUUG AUCAA		NM_001256054.2_2163 -2185_C1A_as	2163-2185
AD-349454.1	UGGGAUUCAGUCUG UAGAAAU	NM_001256054.2_2 224-2244_s	2224-2244	AUUUCUACAGACUGAAUC CCAGG		NM_001256054.2_2222 -2244_as	2222-2244
AD-349474.1	UGUCUAAUAGUUCUC UAUAGU	NM_001256054.2_2 244-2264_s	2244-2264	ACUAUAGAGAACUAUUAG ACAUU		NM_001256054.2_2242 -2264_as	2242-2264
AD-349477.1	CUAAUAGUUCUCUAU AGUCCU	NM_001256054.2_2 247-2267_s	2247-2267	AGGACUAUAGAGAACUAU UAGAC		NM_001256054.2_2245 -2267_as	2245-2267
AD-349610.1	AGCCAAAUUGAAAUG UGCACU	NM_001256054.2_2 419-2439_C21U_s	2419-2439	AGUGCACAUUUCAAUUUG GCUCA		NM_001256054.2_2417 -2439_G1A_as	2417-2439
AD-349692.1	UUCUUGCUAAGUCUU ACCAUU	NM_001256054.2_2 524-2544_G21U_s	2524-2544	AAUGGUAAGACUUAGCAA GAAGA		NM_001256054.2_2522 -2544_C1A_as	2522-2544
AD-349847.1	UGCAAUAGGCUAUAA GGAAUA	NM_001256054.2_2 699-2719_s	2699-2719	UAUUCCTUAUAGCCUAUU GCAGG		NM_001256054.2_2697 -2719_as	2697-2719
AD-349848.1	GCAAUAGGCUAUAAG GAAUAU	NM_001256054.2_2 700-2720_G21U_s	2700-2720	AUAUUCCUUAUAGCCUAU UGCAG		NM_001256054.2_2698 -2720_C1A_as	2698-2720

[1039]

AD-350108.1	GAAAGGUCAUAAAUAG CUUUCU	NM_001256054.2_3013-3033_C21U_s	3013-3033	AGAAAGCUAUUAUGACCU UUCAC	NM_001256054.2_3011-3033_G1A_as	3011-3033
AD-350109.1	AAAGGUCAUAAUAGC UUUCCU	NM_001256054.2_3014-3034_C21U_s	3014-3034	AGGAAAGCUAUUAUGACC UUUCA	NM_001256054.2_3012-3034_G1A_as	3012-3034
AD-350179.1	AACUAGAUGACUGUU GUACUU	NM_001256054.2_3084-3104_G21U_s	3084-3104	AAGUACAACAGUCAUCUA GUUCA	NM_001256054.2_3082-3104_C1A_as	3082-3104
AD-350197.1	CUGUAGCUCAGUCAU UUAAAA	NM_001256054.2_3102-3122_s	3102-3122	UUUUAAAUGACUGAGCUA CAGUA	NM_001256054.2_3100-3122_as	3100-3122
AD-350270.1	CUGGUUUAUUGUAC UGUUAUA	NM_001256054.2_3211-3231_s	3211-3231	UAUAACAGUACAAUAAAC CAGCC	NM_001256054.2_3209-3231_as	3209-3231
AD-350329.1	AUGUGUAAACAUUGU UAUAUA	NM_001256054.2_3270-3290_s	3270-3290	UAUAUAACAAUGUUUACA CAUGC	NM_001256054.2_3268-3290_as	3268-3290

[1040]

表 13. 靶向 C9orf72 的编码区的 dsRNA 剂的修饰正义及反义链序列

双链体 ID	正义序列 5'至 3'	SEQ ID NO:	反义序列 5'至 3'	SEQ ID NO:	mRNA 靶标序列 5'至 3'	SEQ ID NO:
AD-347430.1	usgscgucAfaAfCfAfgcgacaaguuL96		asAfscuug(Tgn)cgcuguUfuGfacgcascsc		GGUGCGUCAAACAGCGACAAGUU	
AD-347432.1	csgsucaaAfcAfGfCfgacaaguucuL96		asGfsaacu(Tgn)gucgcuGfuUfugacgscsa		UGCGUCAAACAGCGACAAGUUCC	
AD-347454.1	cscscacgUfaAfAfAfgaugacgcuuL96		asAfsgcgcu(Cgn)aucuuuUfaCfgugggscsg		CGCCCACGUAAAAGAUGACGCUU	
AD-347455.1	cscsacguAfaAfAfAfGfaugacgcuuuL96		asAfsagcg(Tgn)caucuuUfuAfcguggsgsc		GCCCACGUAAAAGAUGACGCUUG	
AD-347570.1	asgsugauGfuCfGfAfcucuuugccuL96		asGfsgcaa(Agn)gagucgAfcAfucacusgsc		GCAGUGAUGUCGACUCUUUGCCC	
AD-347602.1	csasgcugUfuGfCfCfaagacagagaL96		usCfsucug(Tgn)cuuggcAfaCfagcugsgsa		UCCAGCUGUUGCCAAGACAGAGA	
AD-347603.1	asgscuguUfgCfCfAfagacagagauL96		asUfscucu(Ggn)ucuuggCfaAfcagcusgsg		CCAGCUGUUGCCAAGACAGAGAU	
AD-347606.1	usgsuugcCfaAfGfAfcagagauuguL96		asCfsauuc(Tgn)cugucuUfgGfcaacasgsc		GCUGUUGCCAAGACAGAGAUUGC	
AD-347610.1	gscscaagAfcAfGfAfgauugcuuuaL96		usAfsaagc(Agn)aucucuGfuCfuuggcsasa		UUGCCAAGACAGAGAUUGCUUUA	
AD-347612.1	csasgacAfgAfGfAfuugcuuuaauL96		asUfsuaaa(Ggn)caaucuCfuGfucuugsgsc		GCCAAGACAGAGAUUGCUUUAAG	
AD-347613.1	asasgacaGfaGfAfUfugcuuuaaguL96		asCfsuuaa(Agn)gcaaucUfcUfgucuusgsg		CCAAGACAGAGAUUGCUUUAAGU	
AD-347614.1	asgsacagAfgAfUfUfgcuuuaaguuL96		asAfscuua(Agn)agcaauCfuCfugucususg		CAAGACAGAGAUUGCUUUAAGUG	
AD-347615.1	gsascagaGfaUfUfGfcuuuaaguguL96		asCfsacuu(Agn)aagcaaUfcUfcugucsusu		AAGACAGAGAUUGCUUUAAGUGG	
AD-347650.1	usasgcagCfuAfCfUfuuugcuuacuL96		asGfsuaag(Cgn)aaaagnAfgCfugcuasasu		AUUAGCAGCUACUUUUGCUUACU	
AD-347674.1	ascsaauaUfuCfUfUfgguccuagauL96		asUfscuag(Ggn)accaagAfaUfauuguscsc		GGACAAUAUUCUUGGUCCUAGAG	
AD-347676.1	asasuauuCfuUfGfGfuccuagaguaL96		usAfscucu(Agn)ggaccaAfgAfauauusgsu		ACAAUAUUCUUGGUCCUAGAGUA	
AD-347677.1	asusauucUfuGfGfUfccuagaguaaL96		usUfsacuc(Tgn)aggaccAfaGfaauausug		CAAUAUUCUUGGUCCUAGAGUAA	
AD-347716.1	asgsacagAfaCfAfGfguacucucaL96		usGfsagaa(Ggn)uaccugUfuCfugucususu		AAAGACAGAACAGGUACUUCUCA	
AD-347773.1	usasaaugGfaGfAfAfauccuucgaaL96		usUfscgaa(Ggn)gauuucUfcCfauuuasgsa		UCUAAAUGGAGAAAUCCUUCGAA	
AD-347778.1	gsgsagaaAfuCfCfUfucgaaaugcaL96		usGfscauu(Tgn)cgaaggAfuUfucuccsasu		AUGGAGAAAUCCUUCGAAAUGCA	

[1041]

双链体 ID	正义序列 5'至 3'	SEQ ID NO:	反义序列 5'至 3'	SEQ ID NO:	mRNA 靶标序列 5'至 3'	SEQ ID NO:
AD-347842.1	csasuuaaUfcUfUfUfgauggaaacuL96		asGfsuuuc(Cgn)aucaaaGfaUfuaaugsasa		UUCAUUAAUCUUUGAUGGAAACU	
AD-347843.1	asusuaauCfuUfUfGfauggaaacuuL96		asAfsguuu(Cgn)caucaaAfgAfuaaugsga		UCAUUAAUCUUUGAUGGAAACUG	
AD-347863.1	csasuaugGfaCfUfAfucaauuauuuL96		asUfsauaa(Tgn)ugauagUfcCfauaugsusg		CACAUAUGGACUAUCAAUUAUUAC	
AD-347865.1	usasuggaCfuAfUfCfaauuauacuuL96		asAfsguau(Agn)auugauAfgUfccauasusg		CAUAUGGACUAUCAAUUAUUACUU	
AD-347867.1	usgsgacuAfuCfAfAfuuauacuucuL96		asGfsaagu(Agn)uaauugAfuAfguccasusa		UAUGGACUAUCAAUUAUUACUUCC	
AD-347874.1	uscsaauuAfuAfCfUfuccacagacaL96		usGfsucug(Tgn)ggaaguAfuAfauugasusa		UAUCAAUUAUACUUCCACAGACA	
AD-347892.1	ascsagaaCfuUfAfGfuuucuaccuuL96		asAfsggua(Ggn)aaacuaAfgUfucuguscsu		AGACAGAACUUAGUUUCUACCUC	
AD-347893.1	csasgaacUfuAfGfUfuucuaccucuL96		asGfsaggu(Agn)gaaacuAfaGfuucugsusc		GACAGAACUUAGUUUCUACCUCC	
AD-347923.1	gsasguguGfuGfUfUfgauagauuaaL96		usUfsaauc(Tgn)aucaacAfcAfcacucsusa		UAGAGUGUGUUGAUAGAUUAA	
AD-347924.1	asgsugugUfgUfUfGfauagauuaauL96		asUfsuaau(Cgn)uaucaaCfaCfacacuscsu		AGAGUGUGUUGAUAGAUUAAC	
AD-347926.1	usgsugugUfuGfAfUfagauuaacauL96		asUfsguua(Agn)ucuaucAfaCfacacascsu		AGUGUGUUGAUAGAUUAACAC	
AD-347942.1	asascacaUfuUfAfAfucggaaaguL96		asCfsuuuc(Cgn)ggauuaUfaUfguguusasa		UUAACACAUAUAAUCCGGAAAGG	
AD-347993.1	usgsuccaGfaAfGfAfuuauucuuagaL96		usCfsuaag(Agn)uaaucuUfcUfggacasusu		AAUGUCCAGAAGAUUAUCUUAGA	
AD-348087.1	gsasacugCfuUfUfCfaucuaugaaaL96		usUfsucau(Agn)gaugaaAfgCfaguucscsa		UGGAACUGCUUUCAUCUAUGAAA	
AD-348136.1	usasgcugAfuAfCfAfguacucaauuL96		asAfsuuga(Ggn)uacuguAfuCfagcuasusa		UAUAGCUGAUACAGUACUCAAUG	
AD-348140.1	usgsauacAfgUfAfCfucaaugaugaL96		usCfsauca(Tgn)ugaguaCfuGfuaucasgsc		GCUGAUACAGUACUCAAUGAUGA	
AD-348176.1	csusgucaUfgAfAfGfgcuuucuucuL96		asGfsaaga(Agn)agccuuCfaUfgacagcsu		AGCUGUCAUGAAGGCUUUCUUCU	
AD-348231.1	usgsuuccGfuUfGfUfagagguaguuL96		asCfsuacc(Tgn)acuacaAfcGfgaacasgsc		GCUGUUCCGUUGUAGUAGGUAGC	
AD-348341.1	asuscauuUfaAfAfUfuagagucaguL96		asCfsugac(Tgn)cauauuUfaAfauguusgsa		UCAUCAUUUAAAUAUGAGUCAGG	
AD-348377.1	cscsugcuAfaAfGfGfauucaacuguL96		asCfsaguu(Ggn)aauccuUfuAfgcaggscsc		GGCCUGCUAAAGGAUUCAACUGG	
AD-348487.1	csascccuGfuCfAfUfgaacauauuuL96		asAfsauau(Ggn)uucaugAfcAfgggugsgsc		GCCACCCUGUCAUGAACAUAUUU	
AD-348497.1	usgsaacaUfaUfUfUfauaaucagcuL96		asGfscuga(Tgn)uauaaaUfaUfguucasusg		CAUGAACAUAUUUUAUAAUCAGCG	

[1042]

双链体 ID	正义序列 5'至 3'	SEQ ID NO:	反义序列 5'至 3'	SEQ ID NO:	mRNA 靶标序列 5'至 3'	SEQ ID NO:
AD-348500.1	ascsauauUfuAfUfAfaucagcguauL96		asUfsacgc(Tgn)gauuauAfaAfuaugususcc		GAACAUAUUUAUAAUCAGCGUAG	
AD-348502.1	asusauuuAfuAfAfUfcagcguaguaL96		asUfscuac(Ggn)cuguauuAfuAfaauausgsu		ACAUAUUUAUAAUCAGCGUAGAU	
AD-348578.1	uscsaggaUfaCfGfAfucuacuacauL96		asUfsguag(Agn)ugauucgUfaUfccugasgsc		GCUCAGGAUACGAUCAUCUACAC	
AD-348588.1	asuscauuCfaCfAfCfugacgaaaguL96		asCfsuuuc(Ggn)ucagugUfaGfnaugauscsg		CGAUCAUCUACACUGACGAAAGC	
AD-348590.1	csasucuaCfaCfUfGfacgaaagcuuL96		asAfsgcuu(Tgn)cgucagUfgUfagaugsasu		AUCAUCUACACUGACGAAAGCUU	
AD-348591.1	asuscuacAfcUfGfUfAfcgaaagcuuuL96		asAfsagcu(Tgn)ucgucaGfuGfuagausgsa		UCAUCUACACUGACGAAAGCUUU	
AD-348596.1	csascugaCfgAfAfAfAfgcuuuacucuL96		asGfsagua(Agn)agcuuuCfgUfcagugsusa		UACACUGACGAAAGCUUUACUCC	
AD-348597.1	ascsugacGfaAfAfAfGfcuuuacucccuL96		asGfsgagu(Agn)aagcuuUfcGfucagusgsu		ACACUGACGAAAGCUUUACUCCU	
AD-348598.1	csusgacgAfaAfAfGfCfuuuacuccuuL96		asAfsggag(Tgn)aaagcuUfuCfgucagsusg		CACUGACGAAAGCUUUACUCCUG	
AD-348630.1	ususucaaGfaUfGfUfUfcuuacacagaL96		usCfsugug(Tgn)aagacaUfcUfugaaasasa		UUUUUCAAGAUGUCUUACACAGA	
AD-348639.1	gsuscuuaCfaCfAfAfGfagacacucuaL96		usAfsgagu(Ggn)ucucugUfgUfaagacsasu		AUGUCUUACACAGAGACACUCUA	
AD-348672.1	csusggauCfaGfGfUfUfcuuucagcuuL96		asAfsgcug(Agn)aagaccUfgAfuccagsgsa		UCCUGGAUCAGGUCUUUCAGCUG	
AD-348708.1	csuscagaAfgUfAfCfuuuccuugcaL96		usGfscaag(Ggn)aaaguaCfuUfcugagsasg		CUCUCAGAAGUACUUUCCUUGCA	
AD-348734.1	uscsuacuUfgUfCfCfUfucacagaaaL96		usUfsucug(Tgn)gaaggaCfaAfguagasasa		UUUCUACUUGUCCUUCACAGAAA	
AD-348751.1	gsasaagCfcUfUfUfGfacacuauauaaL96		usUfsauua(Ggn)ugucaaGfgCfuuuucsusg		CAGAAAAGCCUUGACACUAAUAA	
AD-348787.1	cscsuuuaAfaUfCfCfUfcuucggaacuL96		asGfsuucc(Ggn)aagagaUfuUfaaggsgsc		GCCCUUUAAAUCUCUUCGGAACC	
AD-348788.1	csusuuaaAfuCfUfCfUfucggaaccuL96		asGfsguuc(Cgn)gaagagAfuUfuaaagsgsg		CCCUUUAAAUCUCUUCGGAACCU	
AD-348789.1	ususuaaaUfcUfCfUfUfucggaaccuuL96		asAfsgguu(Cgn)-gaagaGfaUfuuaaasgsg		CCUUUAAAUCUCUUCGGAACCUG	
AD-348836.1	gsgsgcgaUfcUfUfUfAfacauaauaaaL96		asUfsuauu(Agn)uguuaaGfaUfcgcccsusc		GAGGGCGAUCUUAACAUAAUAAU	
AD-348842.1	uscsuuaaCfaUfAfAfuuauggcucuL96		asGfsagcc(Agn)uuauuaUfgUfuaagasusc		GAUCUUAACAUAAUUAAUGGCUCU	
AD-348895.1	ususaucuUfuGfGfAfagaccuuucuL96		asGfsaaag(Ggn)ucuucc-AfaAfgauaasasa		UUUUAUCUUUGGAAGACCUUUCU	
AD-348904.1	gsasagacCfuUfUfCfuacacuaguuL96		asAfscuag(Tgn)guagaaAfgGfucuucscsa		UGGAAGACCUUUCUACACUAGUG	

[1043]

双链体 ID	正义序列 5'至 3'	SEQ ID NO:	反义序列 5'至 3'	SEQ ID NO:	mRNA 靶标序列 5'至 3'	SEQ ID NO:
AD-348905.1	asasgaccUfuUfCfUfcfacuaguguL96		asCfsacua(Ggn)uguagaAfaGfgucuuscsc		GGAAGACCUUUCUACACUAGUGU	
AD-348906.1	asgsaccuUfuCfUfAfcfacuaguguuL96		asAfscacu(Agn)guguagAfaAfggucuusc		GAAGACCUUUCUACACUAGUGUG	
AD-348926.1	gscsaagaAfcGfAfGfuaguucuaauL96		asUfsuaga(Agn)caucucGfuUfcuugcsasc		GUGCAAGAACGAGAUGUUCUAAUGU	
AD-348930.1	gsasacgaGfaUfGfUfcuaaugacuL96		asGfsucau(Tgn)agaacaUfcUfcgucsusu		AAGAACGAGAUGUUCUAAUGACU	
AD-348931.1	asascgagAfuGfUfCfuaaugacuuL96		asAfsguca(Tgn)uagaacAfuCfucguuscsu		AGAACGAGAUGUUCUAAUGACUU	
AD-348932.1	ascsgagaUfgUfUfCfuaaugacuuuL96		asAfsaguc(Agn)uuagaaCfaUfcucgususc		GAACGAGAUGUUCUAAUGACUUU	
AD-348957.1	asusguguAfaCfUfUfuaaagccuaL96		usAfsggcu(Tgn)auuaagUfuAfcacaususu		AAAUGUGUAACUUAAUAAGCCUA	
AD-348959.1	gsusguaaCfuUfAfAfuaagccuauuL96		asAfsuagg(Cgn)uuauuuaAfgUfuacacsasu		AUGUGUAACUUAAUAAGCCUAUU	
AD-348961.1	gsusaacuUfuAfUfAfagccuauucuL96		asGfsauua(Ggn)gcuaauUfuAfguuacsasc		GUGUAACUUAAUAAGCCUAUUCC	
AD-348962.1	usasacuuAfaUfAfAfgccuauuccaL96		usGfsgaau(Agn)ggcuuaUfuAfaguuascsa		UGUAACUUAAUAAGCCUAUUCCA	
AD-348963.1	asascuuaAfuAfAfAfGfccuauuccauL96		asUfsggaa(Tgn)aggcuuAfuUfaaguusasc		GUAACUUAAUAAGCCUAUUCCAU	
AD-349042.1	gsusuaagUfaAfGfUfuacacuacauL96		asUfsguag(Tgn)guaacuUfaCfuuaacsusg		CAGUUAAGUAAGUUACACUACAG	
AD-349048.1	usasaguuAfcAfCfUfcfaguucucaL96		usGfsagaa(Cgn)uguaguGfuAfacuuascsu		AGUAAGUUACACUACAGUUCUCA	
AD-349051.1	gsusuacaCfuAfCfAfguucucacaaL96		usUfsguga(Ggn)aacuguAfgUfguaacsusu		AAGUUACACUACAGUUCUCACAA	
AD-349126.1	csasgaccUfuUfGfUfuuacaauauaL96		usAfsuauu(Ggn)uaaacaUfuGfgucugsusa		UACAGACCUAUGUUUACAAUAUA	
AD-349127.1	asgsaccuAfuGfUfUfuacaauauaaL96		usUfsauau(Tgn)guaaacAfuAfggucusgsu		ACAGACCUAUGUUUACAAUAUAA	
AD-349395.1	gsasucaaGfcAfGfAfuguuuaauuuL96		asAfsauua(Agn)acaucuGfcUfugaucsasa		UUGAUCAAGCAGAUGUUUAAUUG	
AD-349454.1	usgsggauUfcAfGfUfCfuguagaaauL96		asUfsuucu(Agn)cagacuGfaAfucccasgsg		CCUGGGAUUCAGUCUGUAGAAAU	
AD-349474.1	usgsucuaAfuAfGfUfUfcucuauaguL96		asCfsuaua(Ggn)agaacuAfuUfagacasusu		AAUGUCUAAUAGUUCUCUAUAGU	
AD-349477.1	csusaauaGfuUfCfUfCfuauaguccuL96		asGfsgacu(Agn)uagagaAfcUfauuagsasc		GUCUAAUAGUUCUCUAUAGUCCU	
AD-349610.1	asgsccaaAfuUfGfUfAfaaugugcacuL96		asGfsugca(Cgn)auuucaAfuUfuggcuscsa		UGAGCCAAAUUGAAAUGUGCACC	
AD-349692.1	ususcuugCfuAfAfGfUfcuuaccauuuL96		asAfsuggu(Agn)agacuuAfgCfaagaasgsa		UCUUCUUGCUAAGUCUUACCAUG	

[1044]

双链体 ID	正义序列 5'至 3'	SEQ ID NO:	反义序列 5'至 3'	SEQ ID NO:	mRNA 靶标序列 5'至 3'	SEQ ID NO:
AD-349847.1	usgscaauAfgGfCfUfauaaggaauaL96		usAfsuucc(Tgn)uauagcCfuAfuugcasgsg		CCUGCAAUAGGCUAUAAGGAAUA	
AD-349848.1	gscsauaGfgCfUfAfuaaggauauaL96		asUfsauuc(Cgn)uauagCfcUfauugcsasg		CUGCAAUAGGCUAUAAGGAAUAG	
AD-350108.1	gsasaggUfcAfUfAfauagcuuucuL96		asGfsaaag(Cgn)uauuauGfaCfcuuucsasc		GUGAAAGGUCAUAAUAGCUUUCC	
AD-350109.1	asasagguCfaUfAfAfuagcuuuccuL96		asGfsgaaa(Ggn)cuauuaUfgAfccuuuscsa		UGAAAGGUCAUAAUAGCUUUCCC	
AD-350179.1	asascuagAfuGfAfCfuguguacuuL96		asAfsguac(Agn)acagucAfuCfuaguuscsa		UGAACUAGAUGACUGUUGUACUG	
AD-350197.1	csusguagCfuCfAfGfucauuuaaaL96		usUfsuuaa(Agn)ugacugAfgCfuacagsusa		UACUGUAGCUCAGUCAUUUAAAA	
AD-350270.1	csusgguUfaUfUfGfuacuguauaL96		usAfsuaac(Agn)guacaaUfaAfaccagscsc		GGCUGGUUUAUUGUACUGUUAUA	
AD-350329.1	asusguguAfaAfCfAfuuguauauaL96		usAfsuaua(Agn)caauguUfuAfcacausgsc		GCAUGUGUAAACAUUGUUAUAUA	

[1045] 表14.靶向外显子1a及1b之间的内含子的dsRNA剂的UC9orf72未修饰正义及反义链序列

[1046]

双链体 ID	正义序列 5’ 至 3’	SEQ ID NO:	反义序列 5’ 至 3’	SEQ ID NO:
AD-463863.1	AAAGACCUGAUAAAGAUUAAU		AUUAAUCUUUAUCAGGUCUUUUC	
AD-463862.1	AAAAGACCUGAUAAAGAUUAA		UUAAUCUUUAUCAGGUCUUUUCU	
AD-463869.1	CUGAUAAAGAUUAACCAGAAU		AUUCUGGUUAAUCUUUAUCAGGU	
AD-463873.1	AAAGAUUAACCAGAAGAAAAU		AUUUUCUUCUGGUUAAUCUUUAU	
AD-463872.1	AUAAAGAUUAACCAGAAGAAA		UUUCUUCUGGUUAAUCUUUAUCA	
AD-463860.1	AGAAAAGACCUGAUAAAGAUU		AAUCUUUAUCAGGUCUUUUCUUG	
AD-463847.1	GCUGAGGGUGAACAAGAAAAU		AUUUUCUUGUUCACCCUCAGCGA	
AD-463871.1	GAUAAAGAUUAACCAGAAGAA		UUCUUCUGGUUAAUCUUUAUCAG	
AD-463878.1	AACCAGAAGAAAACAAGGAGU		ACUCCUUGUUUUCUUCUGGUUAA	
AD-463875.1	AUUAACCAGAAGAAAACAAGU		ACUUGUUUUCUUCUGGUUAAUCU	
AD-463877.1	UAACCAGAAGAAAACAAGGAU		AUCCUUGUUUUCUUCUGGUUAAU	
AD-463857.1	AACAAGAAAAGACCUGAUAAA		UUUAUCAGGUCUUUUCUUGUUCA	
AD-463861.1	GAAAAGACCUGAUAAAGAUUA		UAAUCUUUAUCAGGUCUUUUCUU	
AD-463858.1	ACAAGAAAAGACCUGAUAAAU		AUUUAUCAGGUCUUUUCUUGUUC	
AD-463864.1	AAGACCUGAUAAAGAUUAACU		AGUUAAUCUUUAUCAGGUCUUUU	
AD-463866.1	GACCUGAUAAAGAUUAACCAU		AUGGUUAAUCUUUAUCAGGUCUU	

[1047]

AD-463870.1	UGAUAAAGAUUAACCAGAAGA		UCUUCUGGUUAAUCUUUAUCAGG	
AD-463867.1	ACCUGAUAAAGAUUAACCAGA		UCUGGUUAAUCUUUAUCAGGUCU	
AD-463859.1	AAGAAAAGACCUGAUAAAGAU		AUCUUUAUCAGGUCUUUUCUUGU	
AD-463848.1	CUGAGGGUGAACAAGAAAAGA		UCUUUUCUUGUUCACCCUCAGCG	
AD-463855.1	UGAACAAGAAAAGACCUGAUA		UAUCAGGUCUUUUCUUGUUCACC	
AD-463850.1	GAGGGUGAACAAGAAAAGACU		AGUCUUUUCUUGUUCACCCUCAG	
AD-463854.1	GUGAACAAGAAAAGACCUGAU		AUCAGGUCUUUUCUUGUUCACCC	
AD-463874.1	AAGAUUAACCAGAAGAAAACA		UGUUUUCUUCUGGUUAAUCUUUA	
AD-463851.1	AGGGUGAACAAGAAAAGACCU		AGGUCUUUUCUUGUUCACCCUCA	
AD-463868.1	CCUGAUAAAGAUUAACCAGAA		UUCUGGUUAAUCUUUAUCAGGUC	
AD-463832.1	UGCUCUCACAGUACUCGCUGA		UCAGCGAGUACUGUGAGAGCAAG	
AD-463853.1	GGUGAACAAGAAAAGACCUGA		UCAGGUCUUUUCUUGUUCACCCU	
AD-463844.1	CUCGCUGAGGGUGAACAAGAA		UUCUUGUUCACCCUCAGCGAGUA	
AD-463846.1	CGCUGAGGGUGAACAAGAAAA		UUUUCUUGUUCACCCUCAGCGAG	
AD-463839.1	CAGUACUCGCUGAGGGUGAAU		AUUCACCCUCAGCGAGUACUGUG	
AD-463845.1	UCGCUGAGGGUGAACAAGAAA		UUUCUUGUUCACCCUCAGCGAGU	
AD-463820.1	CUCCCCACUACUUGCUCUCAU		AUGAGAGCAAGUAGUGGGGAGAG	
AD-463884.1	ACAACCGCAGCCUGUAGCAAU		AUUGCUACAGGCUGCGGUUGUUU	
AD-463831.1	UUGCUCUCACAGUACUCGCUU		AAGCGAGUACUGUGAGAGCAAGU	

[1048]

AD-463842.1	UACUCGCUGAGGGUGAACAAU		AUUGUUCACCCUCAGCGAGUACU	
AD-463893.1	GUAGCAAGCUCUGGAACUCAU		AUGAGUUCCAGAGCUUGCUACAG	
AD-463899.1	AGCUCUGGAACUCAGGAGUCU		AGACUCCUGAGUUCCAGAGCUUG	
AD-463822.1	CCCCACUACUUGCUCUCACAU		AUGUGAGAGCAAGUAGUGGGGAG	
AD-463834.1	CUCUCACAGUACUCGCUGAGU		ACUCAGCGAGUACUGUGAGAGCA	
AD-463815.1	UGUGUUUUUGUUUUUCCCACU		AGUGGGAAAAACAAAAAACACACA	
AD-463856.1	GAACAAGAAAAGACCUGAUAA		UUAUCAGGUCUUUUCUUGUUCAC	
AD-463879.1	ACCAGAAGAAAACAAGGAGGU		ACCUCCUUGUUUUCUUCUGGUUA	
AD-463894.1	UAGCAAGCUCUGGAACUCAGU		ACUGAGUUCCAGAGCUUGCUACA	
AD-463865.1	AGACCUGAUAAAGAUUAACCA		UGGUUAAUCUUUAUCAGGUCUUU	
AD-463841.1	GUACUCGCUGAGGGUGAACAA		UUGUUCACCCUCAGCGAGUACUG	
AD-463829.1	ACUUGCUCUCACAGUACUCGU		ACGAGUACUGUGAGAGCAAGUAG	
AD-463889.1	GCCUGUAGCAAGCUCUGGAAU		AUUCCAGAGCUUGCUACAGGCUG	
AD-463897.1	CAAGCUCUGGAACUCAGGAGU		ACUCCUGAGUUCCAGAGCUUGCU	
AD-463814.1	GUGUGUUUUUGUUUUUCCCAU		AUGGGAAAAACAAAAACACACAC	
AD-463825.1	CACUACUUGCUCUCACAGUAU		AUACUGUGAGAGCAAGUAGUGGG	
AD-463898.1	AAGCUCUGGAACUCAGGAGUU		AACUCCUGAGUUCCAGAGCUUGC	
AD-463827.1	CUACUUGCUCUCACAGUACUU		AAGUACUGUGAGAGCAAGUAGUG	
AD-463828.1	UACUUGCUCUCACAGUACUCU		AGAGUACUGUGAGAGCAAGUAGU	

[1049]

AD-463885.1	CAACCGCAGCCUGUAGCAAGU		ACUUGCUACAGGCUGCGGUUGUU	
AD-463852.1	GGGUGAACAAGAAAAGACCUU		AAGGUCUUUUCUUGUUCACCCUC	
AD-463830.1	CUUGCUCUCACAGUACUCGCU		AGCGAGUACUGUGAGAGCAAGUA	
AD-463876.1	UUAACCAGAAGAAAACAAGGA		UCCUUGUUUUCUUCUGGUUAAUC	
AD-463819.1	UCUCCCCACUACUUGCUCUCA		UGAGAGCAAGUAGUGGGGAGAGA	
AD-463833.1	GCUCUCACAGUACUCGCUGAU		AUCAGCGAGUACUGUGAGAGCAA	
AD-463900.1	GCUCUGGAACUCAGGAGUCGU		ACGACUCCUGAGUUCCAGAGCUU	
AD-463838.1	ACAGUACUCGCUGAGGGUGAA		UUCACCCUCAGCGAGUACUGUGA	
AD-463886.1	AACCGCAGCCUGUAGCAAGCU		AGCUUGCUACAGGCUGCGGUUGU	
AD-463821.1	UCCCCACUACUUGCUCUCACA		UGUGAGAGCAAGUAGUGGGGAGA	
AD-463826.1	ACUACUUGCUCUCACAGUACU		AGUACUGUGAGAGCAAGUAGUGG	
AD-463881.1	GGAAACAACCGCAGCCUGUAU		AUACAGGCUGCGGUUGUUUCCCU	
AD-463890.1	CCUGUAGCAAGCUCUGGAACU		AGUUCCAGAGCUUGCUACAGGCU	
AD-463849.1	UGAGGGUGAACAAGAAAAGAU		AUCUUUUCUUGUUCACCCUCAGC	
AD-463892.1	UGUAGCAAGCUCUGGAACUCA		UGAGUUCCAGAGCUUGCUACAGG	
AD-463840.1	AGUACUCGCUGAGGGUGAACA		UGUUCACCCUCAGCGAGUACUGU	
AD-463823.1	CCCACUACUUGCUCUCACAGU		ACUGUGAGAGCAAGUAGUGGGGA	
AD-463888.1	AGCCUGUAGCAAGCUCUGGAA		UUCCAGAGCUUGCUACAGGCUGC	
AD-463835.1	CUCACAGUACUCGCUGAGGGU		ACCCUCAGCGAGUACUGUGAGAG	

[1050]

AD-463824.1	CCACUACUUGCUCUCACAGUA		UACUGUGAGAGCAAGUAGUGGGG	
AD-463896.1	GCAAGCUCUGGAACUCAGGAU		AUCCUGAGUUCCAGAGCUUGCUA	
AD-463891.1	CUGUAGCAAGCUCUGGAACUU		AAGUUCCAGAGCUUGCUACAGGC	
AD-463887.1	CCGCAGCCUGUAGCAAGCUCU		AGAGCUUGCUACAGGCUGCGGUU	
AD-463843.1	ACUCGCUGAGGGUGAACAAGA		UCUUGUUCACCCUCAGCGAGUAC	
AD-463901.1	UCAGGAGUCGCGCGCUAGGGU		ACCCUAGCGCGCGACUCCUGAGU	
AD-463895.1	AGCAAGCUCUGGAACUCAGGA		UCCUGAGUUCCAGAGCUUGCUAC	
AD-463817.1	GUUUUUGUUUUUCCCACCCUU		AAGGGUGGGAAAAACAAAAACAC	
AD-463816.1	GUGUUUUUGUUUUUCCCACCU		AGGUGGGAAAAACAAAAACACAC	
AD-463883.1	AACAACCGCAGCCUGUAGCAA		UUGCUACAGGCUGCGGUUGUUUC	
AD-463902.1	CAGGAGUCGCGCGCUAGGGGU		ACCCCUAGCGCGCGACUCCUGAG	
AD-463904.1	GGAGUCGCGCGCUAGGGGCCU		AGGCCCCUAGCGCGCGACUCCUG	
AD-463837.1	CACAGUACUCGCUGAGGGUGA		UCACCCUCAGCGAGUACUGUGAG	
AD-463880.1	GGGAAACAACCGCAGCCUGUA		UACAGGCUGCGGUUGUUUCCCUC	
AD-463903.1	AGGAGUCGCGCGCUAGGGGCU		AGCCCCUAGCGCGCGACUCCUGA	
AD-463882.1	AAACAACCGCAGCCUGUAGCA		UGCUACAGGCUGCGGUUGUUUCC	
AD-463905.1	GAGUCGCGCGCUAGGGGCCGU		ACGGCCCCUAGCGCGCGACUCCU	
AD-463906.1	AGUCGCGCGCUAGGGGCCGGU		ACCGGCCCCUAGCGCGCGACUCC	

[1051] 表15.C9orf72 dsRNA剂的经修饰正义及反义链序列

[1052]

双链体 ID	正义序列 5'至 3'	SEQ ID NO:	反义序列 5'至 3'	SEQ ID NO:	SEQ ID NO:
AD-463863.1	asasagacCfuGfAfUfaaagauuaauL96		asUfsuaaUfcUfUfuaucAfgGfucuuususc		
AD-463862.1	asasaagaCfcUfGfAfuaaagauuaaL96		usUfsaauCfuUfUfaucaGfgUfcuuuuscsu		
AD-463869.1	csusgauaAfaGfAfUfuaaccagaauL96		asUfsucuGfgUfUfaaucUfuUfaucagsgsu		
AD-463873.1	asasagauUfaAfCfCfagaagaaaauL96		asUfsuuuCfuUfCfugguUfaAfucuuusasu		
AD-463872.1	asusaaagAfuUfAfAfccagaagaaaL96		usUfsucuUfcUfGfguuaAfuCfuuuauscsa		
AD-463860.1	asgsaaaaGfaCfCfUfgauaaagauuL96		asAfsucuUfuAfUfcaggUfcUfuuucususg		
AD-463847.1	gscsugagGfgUfGfAfacaagaaaauL96		asUfsuuuCfuUfGfuucaCfcCfucagcsgsa		
AD-463871.1	gsasuaaaGfaUfUfAfaccagaagaaL96		usUfscuuCfuGfGfuuaaUfcUfuuaucsasg		
AD-463878.1	asasccagAfaGfAfAfaacaaggaguL96		asCfsuccUfuGfUfuuucUfuCfugguusasa		
AD-463875.1	asusuaacCfaGfAfAfgaaaacaaguL96		asCfsuugUfuUfUfcuucUfgGfuuauscsu		
AD-463877.1	usasaccaGfaAfGfAfaaacaaggauL96		asUfsccuUfgUfUfuucuUfcUfgguuasasu		
AD-463857.1	asascaagAfaAfAfGfaccugauaaaL96		usUfsuauCfaGfGfucuuUfuCfuuguuscsa		
AD-463861.1	gsasaaagAfcCfUfGfauaaagauuaL96		usAfsaucUfuUfAfucagGfuCfuuuucsusu		
AD-463858.1	ascsaagaAfaAfGfAfccugauaaauL96		asUfsuuaUfcAfGfgucuUfuUfcuugususc		
AD-463864.1	asasgaccUfgAfUfAfaagauuaacuL96		asGfsuuaAfuCfUfuuauCfaGfgucuususu		
AD-463866.1	gsasccugAfuAfAfAfgauuaaccauL96		asUfsgguUfaAfUfcuuuAfuCfaggucsusu		
AD-463870.1	usgsauaaAfgAfUfUfaaccagaagaL96		usCfsuucUfgGfUfuaauCfuUfuaucasgsg		
AD-463867.1	ascscugaUfaAfAfGfauuaaccagaL96		usCfsuggUfuAfAfucuuUfaUfcaguscsu		

[1053]

AD-463859.1	asasgaaaAfgAfCfCfugauaaagauL96		asUfscuuUfaUfCfagguCfuUfuucuusgsu		
AD-463848.1	csusgaggGfuGfAfAfcaagaaaagaaL96		usCfsuuuUfcUfUfguucAfcCfcucagscsg		
AD-463855.1	usgsaacaAfgAfAfAfagaccugauaaL96		usAfsucaGfgUfCfuuuuCfuUfguucascsc		
AD-463850.1	gsasggguGfaAfCfAfagaaaagacuL96		asGfsucuUfuUfCfuuguUfcAfcccucsasg		
AD-463854.1	gsusgaacAfaGfAfAfaagaccugauL96		asUfscagGfuCfUfuuucUfuGfuucacscsc		
AD-463874.1	asasgauuAfaCfCfAfgaagaaaacaL96		usGfsuuuUfcUfUfcuggUfuAfaucuususa		
AD-463851.1	asgsggugAfaCfAfAfgaaaagaccuL96		asGfsgucUfuUfUfcuugUfuCfacccuscsa		
AD-463868.1	cscsugauAfaAfGfAfuuaaccagaaL96		usUfscugGfuUfAfaucuUfuAfucaggsusc		
AD-463832.1	usgscucuCfaCfAfGfuacucgcugaL96		usCfsagcGfaGfUfacugUfgAfgagcasasg		
AD-463853.1	gsgsugaaCfaAfGfAfaaagaccugaL96		usCfsaggUfcUfUfuucuUfgUfucaccscsu		
AD-463844.1	csuscgcuGfaGfGfGfugaacaagaaL96		usUfscuuGfuUfCfacccUfcAfgcgagsusa		
AD-463846.1	csgscugaGfgGfUfGfaacaagaaaaL96		usUfsuucUfuGfUfucacCfcUfcagcgsasg		
AD-463839.1	csasguacUfcGfCfUfgagggugaauL96		asUfsucaCfcCfUfcagcGfaGfuacugsusg		
AD-463845.1	uscsgcugAfgGfGfUfgaacaagaaL96		usUfsucuUfgUfUfcaccCfuCfagcgasgsu		
AD-463820.1	csusccccAfcUfAfCfuugcucucauL96		asUfsgagAfgCfAfaguaGfuGfgggagsasg		
AD-463884.1	ascsaaccGfcAfGfCfcuguagcaauL96		asUfsugcUfaCfAfggcuGfcGfguugususu		
AD-463831.1	ususgcucUfcAfCfAfguacucgcuuL96		asAfsgcgAfgUfAfcuguGfaGfagcaasgsu		
AD-463842.1	usascucgCfuGfAfGfggugaacaauL96		asUfsuguUfcAfCfccucAfgCfgaguascsu		
AD-463893.1	gsusagcaAfgCfUfCfuggaacucauL96		asUfsgagUfuCfCfagagCfuUfgcuacsasg		

[1054]

AD-463899.1	asgscucuGfgAfAfCfucaggaguc uL96		asGfsacuCfcUfGfaguuCfcAfgagc ususg		
AD-463822.1	cscsccacUfaCfUfUfgcucucaca uL96		asUfsgugAfgAfGfcaagUfaGfugg ggsasg		
AD-463834.1	csuscucaCfaGfUfAfcucgcugag uL96		asCfsucaGfcGfAfguacUfgUfgaga gscsa		
AD-463815.1	usgsuguuUfuUfGfUfuuuuccca cuL96		asGfsuggGfaAfAfaacaAfaAfacac ascsa		
AD-463856.1	gsasacaaGfaAfAfAfgaccugaua aaL96		usUfsaucAfgGfUfcuuuUfcUfugu ucsasc		
AD-463879.1	ascscagaAfgAfAfAfacaaggagg uL96		asCfscucCfuUfGfuuuuCfuUfcug gususa		
AD-463894.1	usasgcaaGfcUfCfUfggaacucag uL96		asCfsugaGfuUfCfcagaGfcUfugcu ascsa		
AD-463865.1	asgsaccuGfaUfAfAfagauuaacc aL96		usGfsguuAfaUfCfuuuaUfcAfggu cususu		
AD-463841.1	gsusacucGfcUfGfAfgggugaaca aL96		usUfsguuCfaCfCfcucaGfcGfagua csusg		
AD-463829.1	ascsuugcUfcUfCfAfcaguacucg uL96		asCfsgagUfaCfUfgugaGfaGfcaag usasg		
AD-463889.1	gscscuguAfgCfAfAfgcucuggaa uL96		asUfsuccAfgAfGfcuugCfuAfcag gcsusg		
AD-463897.1	csasagcuCfuGfGfAfacucaggag uL96		asCfsuccUfgAfGfuuccAfgAfgcu ugscsu		
AD-463814.1	gsusguguUfuUfUfGfuuuuuccc auL96		asUfsgggAfaAfAfacaaAfaAfcaca csasc		
AD-463825.1	csascuacUfuGfCfUfcucacagua uL96		asUfsacuGfuGfAfgagcAfaGfuag ugsgsg		
AD-463898.1	asasgcucUfgGfAfAfcucaggagu uL96		asAfscucCfuGfAfguucCfaGfagcu usgsc		
AD-463827.1	csusacuuGfcUfCfUfcacaguacu uL96		asAfsguaCfuGfUfgagaGfcAfagu agsusg		
AD-463828.1	usascuugCfuCfUfCfacaguacuc uL96		asGfsaguAfcUfGfugagAfgCfaag uasgsu		
AD-463885.1	csasaccgCfaGfCfCfuguagcaag uL96		asCfsuugCfuAfCfaggcUfgCfggu ugsusu		
AD-463852.1	gsgsgugaAfcAfAfGfaaaagaccu uL96		asAfsgguCfuUfUfucuuGfuUfcac ccsusc		

[1055]

AD-463830.1	csusugcuCfuCfAfCfaguacucgcuL96		asGfscgaGfuAfCfugugAfgAfgcaagsusa		
AD-463876.1	ususaaccAfgAfAfGfaaaacaaggaL96		usCfscuuGfuUfUfucuuCfuGfguuaasusc		
AD-463819.1	uscsucccCfaCfUfAfcuugcucucaL96		usGfsagaGfcAfAfguagUfgGfggagasgsa		
AD-463833.1	gscsucucAfcAfGfUfacucgcugauL96		asUfscagCfgAfGfuacuGfuGfagagcsasa		
AD-463900.1	gscsucugGfaAfCfUfcaggagucguL96		asCfsgacUfcCfUfgaguUfcCfagagcsusu		
AD-463838.1	ascsaguaCfuCfGfCfugaggugaaL96		usUfscacCfcUfCfagcgAfgUfacugusgsa		
AD-463886.1	asasccgcAfgCfCfUfguagcaagcuL96		asGfscuuGfcUfAfcaggCfuGfcgguusgsu		
AD-463821.1	uscscccaCfuAfCfUfugcucucacaL96		usGfsugaGfaGfCfaaguAfgUfggggasgsa		
AD-463826.1	ascsuacuUfgCfUfCfucacaguacuL96		asGfsuacUfgUfGfagagCfaAfguagusgsg		
AD-463881.1	gsgsaaacAfaCfCfGfcagccuguauL96		asUfsacaGfgCfUfgcggUfuGfuuuccscsu		
AD-463890.1	cscsuguaGfcAfAfGfcucuggaacuL96		asGfsuucCfaGfAfgcuuGfcUfacaggscsu		
AD-463849.1	usgsagggUfgAfAfCfaagaaaagauL96		asUfscuuUfuCfUfuguuCfaCfccucasgsc		
AD-463892.1	usgsuagcAfaGfCfUfcuggaacucaL96		usGfsaguUfcCfAfgagcUfuGfcuacasgsg		
AD-463840.1	asgsuacuCfgCfUfGfagggugaacaL96		usGfsuucAfcCfCfucagCfgAfguacusgsu		
AD-463823.1	cscscacuAfcUfUfGfcucucacaguL96		asCfsuguGfaGfAfgcaaGfuAfgugggsgsa		
AD-463888.1	asgsccugUfaGfCfAfagcucuggaaL96		usUfsccaGfaGfCfuugcUfaCfaggcusgsc		
AD-463835.1	csuscacaGfuAfCfUfcgcugagggaL96		asCfsccuCfaGfCfgaguAfcUfgugagsasg		
AD-463824.1	cscsacuaCfuUfGfCfucucacaguaL96		usAfscugUfgAfGfagcaAfgUfaguggsgsg		
AD-463896.1	gscsaagcUfcUfGfGfaacucaggauL96		asUfsccuGfaGfUfuccaGfaGfcuugcsusa		

[1056]

AD-463891.1	csusguagCfaAfGfCfucuggaacuuL96		asAfsguuCfcAfGfagcuUfgCfuacagsgsc		
AD-463887.1	cscsgcagCfcUfGfUfagcaagcucuL96		asGfsagcUfuGfCfuacaGfgCfugcggsusu		
AD-463843.1	ascsucgcUfgAfGfGfgugaacaagaL96		usCfsuugUfuCfAfcccuCfaGfcgagusasc		
AD-463901.1	uscsaggaGfuCfGfCfgcgcuaggguL96		asCfsccuAfgCfGfcgcgAfcUfccuggasgsu		
AD-463895.1	asgscaagCfuCfUfGfgaacucaggaL96		usCfscugAfgUfUfccagAfgCfuugcusasc		
AD-463817.1	gsusuuuuGfuUfUfUfucccacccuuL96		asAfsgggUfgGfGfaaaaAfcAfaaaacsasc		
AD-463816.1	gsusguuuUfuGfUfUfuuucccaccuL96		asGfsgugGfgAfAfaaacAfaAfaacacsasc		
AD-463883.1	asascaacCfgCfAfGfccuguagcaaL96		usUfsgcuAfcAfGfgcugCfgGfuuguususc		
AD-463902.1	csasggagUfcGfCfGfcgcuagggguL96		asCfscccUfaGfCfgcgcGfaCfuccugsasg		
AD-463904.1	gsgsagucGfcGfCfGfcuaggggccuL96		asGfsgccCfcUfAfgcgcGfcGfacuccsusg		
AD-463837.1	csascaguAfcUfCfGfcugagggugaL96		usCfsaccCfuCfAfgcgaGfuAfcugugsasg		
AD-463880.1	gsgsgaaaCfaAfCfCfgcagccuguaL96		usAfscagGfcUfGfcgguUfgUfuuccsusc		
AD-463903.1	asgsgaguCfgCfGfCfgcuaggggcuL96		asGfscccCfuAfGfcgcgCfgAfcuccusgsa		
AD-463882.1	asasacaaCfcGfCfAfgccuguagcaL96		usGfscuaCfaGfGfcugcGfgUfuguuuscsc		
AD-463905.1	gsasgucgCfgCfGfCfuaggggccguL96		asCfsggcCfcCfUfagcgCfgCfgacucscsu		
AD-463906.1	asgsucgcGfcGfCfUfaggggccgguL96		asCfscggCfcCfCfuagcGfcGfcgacuscsc		

[1057] 表16.Neuro2a细胞中的C9orf72单一剂量筛选

[1058]

	10 nM 剂量		0.1 nM 剂量	
双链体	Avg % C9orf72 mRNA 保留	SD	Avg % C9orf72 mRNA 保留	SD

[1059]

	10 nM 剂量		0.1 nM 剂量	
双链体	Avg % C9orf72 mRNA 保留	SD	Avg % C9orf72 mRNA 保留	SD
AD-347430.1	108.1586	9.776762	102.926461	9.844402
AD-347432.1	80.82998	7.685495	93.4467535	3.707112
AD-347454.1	101.7639	6.31874	96.0559772	3.124197
AD-347455.1	93.96188	12.4245	103.672652	14.9392
AD-347570.1	100.9925	14.11266	97.2382292	3.232345
AD-347602.1	50.61606	2.769283	90.9130248	14.80726
AD-347603.1	36.75218	3.59788	91.2130671	9.217747
AD-347606.1	14.03328	0.575122	82.1811557	11.31506
AD-347610.1	16.84431	3.344733	84.438968	18.13649
AD-347612.1	11.30135	1.222139	64.3047238	9.535823
AD-347613.1	16.45266	1.644385	87.0981976	4.857487
AD-347614.1	47.69549	7.728849	85.9734577	6.335515
AD-347615.1	40.87763	6.263038	95.9572664	10.7571
AD-347650.1	101.8945	13.09817	94.8255195	10.59428
AD-347674.1	65.11115	6.153852	99.7925784	3.061382
AD-347676.1	55.2216	6.538367	92.3882432	8.410176
AD-347677.1	76.39543	7.73141	98.5375449	9.261759
AD-347716.1	101.7214	5.417192	101.563557	4.413121
AD-347773.1	25.37153	5.786308	90.5692411	9.346695
AD-347778.1	67.86754	25.22616	88.9278585	9.809726
AD-347842.1	83.67195	9.396759	106.321112	12.30094
AD-347843.1	99.14834	12.88462	93.3342087	12.02222
AD-347863.1	43.93743	5.498079	92.8893809	5.933657
AD-347865.1	83.32648	17.19332	112.817387	7.794245
AD-347867.1	69.77927	4.708071	97.6185534	10.16772
AD-347874.1	76.30515	13.31293	111.290499	17.68477
AD-347892.1	96.22058	7.665784	92.4547627	11.2075
AD-347893.1	101.5948	8.86735	105.607076	10.04238
AD-347923.1	82.3263	13.52911	100.608981	17.91672
AD-347924.1	90.86273	9.373496	94.7441667	7.274829
AD-347926.1	82.85491	21.92686	89.7017391	9.976171
AD-347942.1	93.69591	12.95754	91.8512716	17.41955
AD-347993.1	95.05716	13.84628	110.698325	20.13021
AD-348087.1	64.2739	3.190526	98.6940059	13.94537
AD-348136.1	11.22875	0.70978	54.5425144	5.936783
AD-348140.1	20.23686	3.421909	84.7820238	11.42362
AD-348176.1	62.68142	6.009446	97.6143802	19.28814
AD-348231.1	84.50123	10.95544	88.6877429	6.371108
AD-348341.1	96.09791	12.62484	96.4581518	12.15386
AD-348377.1	100.1362	8.605203	110.137549	16.89486

[1060]

	10 nM 剂量		0.1 nM 剂量	
双链体	Avg % C9orf72 mRNA 保留	SD	Avg % C9orf72 mRNA 保留	SD
AD-348487.1	75.55589	10.21049	86.2374386	0.392865
AD-348497.1	82.80696	14.32393	88.5415729	2.818069
AD-348500.1	75.31347	7.550983	95.8111343	11.06454
AD-348502.1	84.82139	14.11956	88.7596585	9.557181
AD-348578.1	70.79848	11.44325	95.3600129	6.623138
AD-348588.1	90.28763	10.19408	94.5206354	6.193889
AD-348590.1	85.64384	22.17813	94.23508	14.42933
AD-348591.1	88.27091	9.957384	97.8838666	13.20964
AD-348596.1	102.3846	15.30875	87.3876578	4.606803
AD-348597.1	90.49934	9.265808	96.2475161	4.061686
AD-348598.1	94.34813	9.170095	106.768924	17.55944
AD-348630.1	78.77815	12.8725	95.328952	8.85476
AD-348639.1	13.70135	2.886124	75.7676877	5.843324
AD-348672.1	89.88466	14.53998	98.09215	18.3019
AD-348708.1	61.84377	5.055461	107.413925	21.38253
AD-348734.1	95.86399	15.87249	98.3209181	10.69515
AD-348751.1	31.79708	4.66811	101.630492	5.905301
AD-348787.1	49.79092	5.894205	86.9442769	22.10655
AD-348788.1	51.1666	11.20645	109.857745	12.05388
AD-348789.1	80.88218	8.107075	99.5021731	10.52775
AD-348836.1	32.20706	2.883623	92.794671	9.647938
AD-348842.1	16.61415	3.691673	98.511271	24.0038
AD-348895.1	71.62264	15.37766	96.3016163	14.16359
AD-348904.1	9.257562	2.22834	53.4658877	12.09395
AD-348905.1	15.74941	1.840235	80.898199	5.782131
AD-348906.1	52.90698	5.482321	89.0445935	4.202231
AD-348926.1	89.17682	4.958206	99.7640306	1.997126
AD-348930.1	20.24304	4.668287	87.9325289	10.38548
AD-348931.1	21.64584	1.151471	86.5235896	8.119257
AD-348932.1	65.4211	6.171016	98.0124995	6.290356
AD-348957.1	99.11614	5.422021	98.130978	17.98878
AD-348959.1	98.72032	18.18334	85.5950571	9.305798
AD-348961.1	95.63409	14.86035	88.1861421	8.042136
AD-348962.1	93.07311	17.28878	96.4886467	20.13191
AD-348963.1	93.0378	6.249871	98.5325806	9.378936
AD-349042.1	87.08801	5.943891	97.8928314	13.51528
AD-349048.1	90.07839	9.309048	97.0782319	9.82998
AD-349051.1	72.66003	12.19094	97.0614884	14.31408
AD-349126.1	102.9632	4.735787	101.144929	11.55466
AD-349127.1	88.62489	11.91672	91.8250081	2.66611

[1061]

	10 nM 剂量		0.1 nM 剂量	
双链体	Avg % C9orf72 mRNA 保留	SD	Avg % C9orf72 mRNA 保留	SD
AD-349395.1	95.46553	16.43203	97.6042107	4.227649
AD-349454.1	88.26711	5.092706	105.115217	3.470468
AD-349474.1	89.97248	8.272942	100.835274	12.23494
AD-349477.1	101.5282	12.84979	106.139698	6.558009
AD-349610.1	103.9541	15.77276	107.645445	12.40821
AD-349692.1	108.9096	12.64303	101.695084	10.31383
AD-349847.1	90.23518	13.76773	88.8762369	3.801429
AD-349848.1	93.67591	6.621208	100.159587	4.107114
AD-350108.1	112.3379	10.06486	98.3414867	5.841237
AD-350109.1	96.98935	9.710059	101.154071	9.850331
AD-350179.1	83.68433	16.73609	102.259228	4.069927
AD-350197.1	83.4485	6.210923	85.9923746	3.243041
AD-350270.1	99.62824	14.44672	90.0604061	6.15077
AD-350329.1	17.11297	2.798394	77.8753892	13.18545

[1062] 表17.BE(2)C细胞中的C9orf72单一剂量筛选

[1063]

	10 nM 剂量		0.1 nM 剂量	
双链体	Avg % C9orf72 mRNA 保留	SD	Avg % C9orf72 mRNA 保留	SD
AD-347430.1	82.88227	13.77676	100.4073	8.711257
AD-347432.1	60.10642	2.72119	70.0206	13.68082
AD-347454.1	110.2161	8.732283	107.3506	1.973967
AD-347455.1	96.54231	10.72406	92.79051	12.60443
AD-347570.1	43.18622	2.33787	98.16339	10.7447
AD-347602.1	60.48774	16.01064	75.41154	5.907048
AD-347603.1	44.44516	5.422459	60.93898	12.51016
AD-347606.1	20.66731	2.206332	43.31785	3.515035
AD-347610.1	43.78379	10.3912	81.08617	2.536212
AD-347612.1	22.41886	2.722727	60.50788	7.288639
AD-347613.1	46.6743	8.170093	65.21552	12.51711
AD-347614.1	72.20537	23.59788	78.17678	20.28306
AD-347615.1	61.48894	5.138911	79.37615	21.31636
AD-347650.1	47.99727	2.522146	104.3003	4.790927
AD-347674.1	80.53007	22.51517	76.06062	17.77269
AD-347676.1	64.25803	19.85368	58.46471	21.36364
AD-347677.1	103.7541	4.87836	89.38045	18.31084
AD-347716.1	39.16332	3.009752	71.40257	4.309919

[1064]

	10 nM 剂量		0.1 nM 剂量	
双链体	Avg % C9orf72 mRNA 保留	SD	Avg % C9orf72 mRNA 保留	SD
AD-347773.1	22.79971	3.801035	82.8196	6.565303
AD-347778.1	55.45244	10.03663	103.55	10.88137
AD-347842.1	85.20458	15.02426	81.30493	11.86813
AD-347843.1	94.05675	18.541	111.0976	13.02574
AD-347863.1	17.14299	1.434336	77.74684	4.954083
AD-347865.1	90.51911	8.09217	96.69731	5.84892
AD-347867.1	37.10558	11.16917	88.67657	15.57378
AD-347874.1	49.26003	4.158059	100.6093	5.739901
AD-347892.1	30.89859	7.317843	71.88334	6.283385
AD-347893.1	30.20136	5.660424	86.78551	13.85252
AD-347923.1	21.31491	3.326146	78.04297	5.51637
AD-347924.1	71.07692	5.337269	99.36227	4.570199
AD-347926.1	27.58301	4.30096	83.90653	5.447661
AD-347942.1	86.30598	10.22092	97.86192	8.429688
AD-347993.1	31.66535	5.027872	85.81969	5.542086
AD-348087.1	39.38533	8.883016	72.86883	3.669532
AD-348136.1	24.30857	5.274965	58.90466	5.373379
AD-348140.1	14.39598	2.368607	67.3472	11.30915
AD-348176.1	78.32646	2.522093	88.55116	10.30097
AD-348231.1	53.4932	4.021274	92.7088	9.539667
AD-348341.1	83.89848	17.44643	86.34048	6.96364
AD-348377.1	50.84853	9.656333	84.3215	11.23426
AD-348487.1	41.47418	6.787419	75.41429	17.5351
AD-348497.1	66.49273	13.48261	86.63238	17.01784
AD-348500.1	19.22007	3.385975	62.65074	10.7548
AD-348502.1	64.79465	11.59712	81.20134	15.65802
AD-348578.1	20.69829	2.038162	65.04362	4.226679
AD-348588.1	22.64718	2.226235	66.52442	6.439306
AD-348590.1	53.82448	9.0131	99.50188	12.42381
AD-348591.1	46.04985	9.076041	72.44525	7.061754
AD-348596.1	36.92846	7.643471	84.48218	9.605091
AD-348597.1	19.59032	4.186554	60.80839	9.97145
AD-348598.1	39.93177	3.173106	80.63123	12.4441
AD-348630.1	71.52849	10.66853	78.44949	9.348965
AD-348639.1	15.6934	2.226095	52.72726	3.204068
AD-348672.1	26.3175	6.176373	61.14771	6.371119
AD-348708.1	35.99064	7.287375	88.86058	10.105
AD-348734.1	31.97164	4.916142	76.4496	8.004687
AD-348751.1	26.67156	2.862387	71.02257	4.224371
AD-348787.1	35.22187	5.595156	75.23651	8.51262

[1065]

	10 nM 剂量		0.1 nM 剂量	
双链体	Avg % C9orf72 mRNA 保留	SD	Avg % C9orf72 mRNA 保留	SD
AD-348788.1	33.69843	3.577929	70.47266	7.116083
AD-348789.1	60.60725	1.970527	110.8428	5.584758
AD-348836.1	67.48432	9.387845	104.4636	8.446368
AD-348842.1	36.89163	3.29892	99.32988	11.88113
AD-348895.1	32.91338	3.743304	89.36052	6.64023
AD-348904.1	28.80843	3.433009	52.81576	9.948805
AD-348905.1	31.06881	5.086494	95.81619	8.079109
AD-348906.1	67.75211	3.326577	104.8576	23.53051
AD-348926.1	81.0301	4.95746	104.3321	11.46042
AD-348930.1	58.13332	18.52752	101.5076	15.73369
AD-348931.1	34.92576	7.411208	74.78969	28.27697
AD-348932.1	27.96506	4.683937	77.32319	10.11088
AD-348957.1	85.18783	9.584405	95.68012	5.275668
AD-348959.1	29.52378	3.036392	71.73005	6.004029
AD-348961.1	44.05841	5.454327	114.1652	14.95256
AD-348962.1	86.72846	11.99514	89.64847	8.637003
AD-348963.1	107.1386	12.93197	107.867	7.059721
AD-349042.1	25.83698	6.208601	73.0493	11.32439
AD-349048.1	51.92857	12.09663	93.208	7.767345
AD-349051.1	32.54422	7.917884	55.10166	12.73971
AD-349126.1	38.02445	8.158304	65.81213	8.34951
AD-349127.1	44.34683	4.329469	61.37972	8.434831
AD-349395.1	38.72672	9.143905	73.90281	6.509425
AD-349454.1	25.74301	1.818085	63.89429	7.087295
AD-349474.1	54.35402	5.785511	104.8446	8.454144
AD-349477.1	35.84411	5.102722	91.68584	6.502661
AD-349610.1	44.55332	5.64699	95.48529	16.9582
AD-349692.1	26.64121	3.550501	71.19351	4.976952
AD-349847.1	42.44005	9.583344	100.7526	8.318353
AD-349848.1	26.44739	2.337312	72.32507	7.431502
AD-350108.1	77.7291	11.47368	110.5336	20.27915
AD-350109.1	26.81206	6.729989	70.7082	8.645896
AD-350179.1	28.94105	4.109557	85.93683	4.82879
AD-350197.1	24.61377	2.991152	48.15037	5.130792
AD-350270.1	34.99259	9.000575	80.03001	11.82625
AD-350329.1	29.75514	4.619727	57.08563	7.22942

[1066] 表18.C9orf72单一剂量筛选

[1067]

	10 nM 剂量	1 nM 剂量	0.1 nM 剂量

[1068]

双链体	Avg % C9orf72 靶标 RNA 保留	SD	Avg % C9orf72 靶标 RNA 保留	SD	Avg % C9orf72 靶标 RNA 保留	SD
AD-463863.1	6.895104	1.887372	17.990786	3.581209	47.74899	9.608242
AD-463862.1	11.71756	1.624716	25.6975889	4.69844	69.58064	11.48016
AD-463869.1	11.89178	1.13751	21.4449784	5.258391	78.90809	4.527617
AD-463873.1	13.12453	1.926293	17.3532715	3.624777	56.40397	14.57048
AD-463872.1	14.31053	2.347295	33.8387941	7.887076	68.98069	9.630744
AD-463860.1	14.38719	4.174403	25.6651884	4.012716	69.34006	8.603515
AD-463847.1	15.94532	2.910257	40.3177154	10.98583	74.72245	3.797427
AD-463871.1	16.59102	3.60679	43.8501782	9.795528	78.33788	11.04167
AD-463878.1	17.44205	6.330127	30.7611339	2.808117	78.51312	2.360064
AD-463875.1	18.06926	5.250411	47.7505591	6.339073	80.28471	4.643024
AD-463877.1	19.76524	2.463983	32.3670028	5.565999	91.14105	4.419216
AD-463857.1	19.92302	5.089927	30.2229624	7.222642	64.72521	12.13353
AD-463861.1	20.76325	3.866795	44.4845299	4.859767	92.14849	6.749253
AD-463858.1	21.82364	7.846073	34.3701191	10.6661	82.0841	8.566596
AD-463864.1	21.8246	3.66252	49.6302605	9.716144	78.01891	3.071723
AD-463866.1	24.0353	4.42879	46.1919626	3.599353	84.54992	5.048781
AD-463870.1	25.46799	4.006128	58.2570405	4.729279	90.10873	4.873256
AD-463867.1	27.63503	6.322693	52.584482	10.57925	93.01992	8.170161
AD-463859.1	27.9419	6.237851	48.112763	4.317608	88.26397	8.121303
AD-463848.1	31.208	7.399623	59.3237856	6.602048	78.3498	9.343049
AD-463855.1	31.6721	3.178719	52.1362368	5.292684	86.86777	6.307716
AD-463850.1	31.72366	9.046134	61.457552	7.846478	89.91207	9.559405
AD-463854.1	33.99302	4.494354	50.3642249	6.049553	98.50773	8.755903
AD-463874.1	34.28805	4.662807	46.9952482	9.765255	81.63979	9.244761
AD-463851.1	36.38047	5.665975	59.8831826	7.846382	88.63486	4.023827
AD-463868.1	37.85565	2.293403	64.9907705	3.362285	86.97633	8.633561
AD-463832.1	40.05936	7.066138	77.1665714	3.454563	88.75029	7.540446
AD-463853.1	40.80135	7.554402	65.5474235	6.99034	87.30056	8.121754
AD-463844.1	41.73267	9.685821	63.364797	9.218938	93.14991	11.03547
AD-463846.1	42.64359	6.597737	62.4440201	5.458658	100.4715	2.361336
AD-463839.1	42.98393	13.02984	68.9062633	2.531826	94.36524	8.202007
AD-463845.1	43.30533	6.617779	67.6577722	3.438657	96.3938	15.75012
AD-463820.1	43.91113	7.978728	69.81747	12.4481	94.38236	4.783487
AD-463884.1	44.73481	8.68839	66.0916778	4.521927	88.30249	4.150882
AD-463831.1	45.28	4.393546	87.4988878	6.794657	96.65482	9.742692
AD-463842.1	47.94548	2.6434	68.3097157	10.49513	89.12627	6.732591
AD-463893.1	48.024	4.50011	70.6327049	14.83107	101.2062	4.477603
AD-463899.1	48.93584	4.533855	79.60743	4.808234	99.56349	7.167818
AD-463822.1	49.22299	11.77449	61.7869785	9.385718	104.7792	5.326017
AD-463834.1	50.54416	7.157611	67.6085094	6.51823	95.15348	6.00028
AD-463815.1	50.7322	8.15209	85.8836494	5.801274	93.306	2.133296

[1069]

	10 nM 剂量		1 nM 剂量		0.1 nM 剂量	
双链体	Avg % C9orf72 靶标 RNA 保留	SD	Avg % C9orf72 靶标 RNA 保留	SD	Avg % C9orf72 靶标 RNA 保留	SD
AD-463856.1	52.09708	9.292538	71.6208025	4.402486	83.66283	7.512359
AD-463879.1	52.29238	12.31902	91.3354901	5.701787	87.11167	8.476857
AD-463894.1	53.83841	5.300078	78.519777	10.16316	98.54044	7.394306
AD-463865.1	53.89827	7.734349	76.9058342	13.81802	90.7827	4.899252
AD-463841.1	55.34068	14.6004	75.3676484	3.249612	86.74545	7.197606
AD-463829.1	55.65741	7.327545	76.05587	5.933388	87.68048	3.095124
AD-463889.1	57.15583	4.218506	79.8912749	10.87685	105.7882	1.681391
AD-463897.1	57.28436	11.24736	78.1815484	3.395568	107.6228	8.576143
AD-463814.1	57.48673	10.52274	76.0543605	6.494248	108.132	15.95589
AD-463825.1	57.72945	9.281364	69.2059288	11.50281	94.58803	7.275608
AD-463898.1	57.81877	7.668013	84.2611527	2.724986	101.0082	5.448219
AD-463827.1	57.94463	7.138324	70.5072119	7.704315	96.08478	2.721961
AD-463828.1	58.4669	6.262051	78.0817792	15.02007	94.57042	6.224673
AD-463885.1	59.80305	6.187644	79.434974	8.667472	102.3	8.49886
AD-463852.1	60.87121	8.593189	85.2019462	7.875974	90.30217	7.204589
AD-463830.1	61.33072	7.241431	88.4420867	4.450489	102.3726	4.326917
AD-463876.1	62.82897	0.59753	89.6575968	7.218984	98.10649	9.501475
AD-463819.1	64.15961	7.129148	85.4054309	15.43181	97.1615	4.038406
AD-463833.1	64.41388	7.336112	82.4725601	7.028475	95.57842	3.279038
AD-463900.1	64.46457	2.572698	83.954137	8.818046	97.58683	11.33522
AD-463838.1	64.65207	5.639705	82.4582348	5.243051	106.8135	10.45547
AD-463886.1	65.06557	11.70007	88.8179271	6.667204	87.75141	6.017168
AD-463821.1	65.36418	10.33992	81.1393994	13.84006	95.46494	7.125634
AD-463826.1	65.52606	15.25068	88.6539653	8.213642	94.64711	10.77663
AD-463881.1	65.66246	7.881975	77.5301467	1.659743	97.81026	5.112844
AD-463890.1	66.23927	5.466641	90.2130612	5.662899	102.2023	6.763745
AD-463849.1	67.84639	19.69218	85.5657586	5.003069	93.71145	5.923616
AD-463892.1	69.13546	9.358767	80.4149188	5.170864	94.32616	7.351974
AD-463840.1	69.45808	7.728181	87.7588532	8.442319	95.12692	10.12238
AD-463823.1	69.87365	14.23887	98.4007125	8.297397	92.44274	8.161986
AD-463888.1	70.24116	9.876654	94.9620706	1.396841	91.30081	3.123677
AD-463835.1	73.50516	17.24816	94.6762699	11.00273	98.88008	1.234605
AD-463824.1	74.94413	8.910127	100.838579	4.251858	87.97097	5.900965
AD-463896.1	75.15816	6.626289	92.2732663	2.956882	91.34906	9.34084
AD-463891.1	75.89363	4.932301	84.3316467	11.76645	101.5825	10.64198
AD-463887.1	77.00152	4.720732	92.3937763	8.106755	107.5829	9.283113
AD-463843.1	77.18108	8.475728	88.0193493	7.350552	101.6992	8.130878
AD-463901.1	79.69214	3.554174	91.3385601	5.151392	99.21389	6.067914
AD-463895.1	84.04935	11.39932	97.3393402	8.076027	99.30656	7.939733

[1070]

	10 nM 剂量		1 nM 剂量		0.1 nM 剂量	
双链体	Avg % C9orf72 靶标 RNA 保留	SD	Avg % C9orf72 靶标 RNA 保留	SD	Avg % C9orf72 靶标 RNA 保留	SD
AD-463817.1	85.08871	13.20444	89.5576574	4.128533	87.36462	5.78227
AD-463816.1	85.70303	9.912715	105.472774	10.27543	104.2863	13.51739
AD-463883.1	85.71968	8.035838	94.1941814	5.718514	92.57683	8.285562
AD-463902.1	86.39055	12.69315	98.0579662	6.445276	95.96132	11.12554
AD-463904.1	87.02583	6.485768	90.3248342	10.19339	92.35205	0.545498
AD-463837.1	87.17962	6.005932	91.36949	4.836236	101.3081	19.91605
AD-463880.1	94.4352	10.15591	93.0563793	7.156635	93.58439	3.276504
AD-463903.1	98.01208	9.460806	92.2505023	13.5983	110.9407	9.669494
AD-463882.1	102.1867	8.188831	98.739041	3.554669	93.32284	9.663378
AD-463905.1	102.8622	17.44574	95.432284	8.163072	109.1775	7.124329
AD-463906.1	104.3926	5.265817	98.6778108	8.880539	103.8829	10.47854

[1071] 实施例6:转基因小鼠体中的体内评估

[1072] 本实施例描述用于在转基因小鼠体内评估C9orf72 RNAi剂的方法,该转基因小鼠表达具有高达例如450个GGGGCC重复序列的人类C9orf72 RNA(参见,例如,Jiang,et al.(supra)。

[1073] 对于实施例5中设计且测定的所选dsRNA剂的能力,评估其降低表达具有高达450个GGGGCC重复序列的人类C9orf72 RNA的小鼠体内的含正义或反义的团簇两者的水平。

[1074] 简而言之,对照同窝小鼠、对于具有高达450个GGGGCC重复序列的人类C9orf72 RNA为杂合的小鼠、以及对于具有高达450个GGGGCC重复序列的人类C9orf72 RNA为纯合的小鼠,经鞘内或经皮下给药单一剂量的感兴趣的dsRNA剂,包括双链体AD-463863.1、AD-463862.1、AD-463869.1、AD-463873.1、AD-463872.1或AD-463860.1,或安慰剂。给药的后两周,将动物处死,收集血液及组织样本,包括大脑皮质、脊髓、肝、脾及颈部淋巴结。

[1075] 如上文讨论的,通过差异性使用转录备选起始位点及终止位点生成了三种C9orf72转录本。因此,为了测定给药dsRNA剂对于有害的含重复序列mRNA的水平的效应,通过qRT-PCR测定皮质及脊髓样本中含重复序列的C9orf 72 mRNA、总C9orf72 mRNA、及含外显子1b的mRNA的水平(参见,例如,上文及Jiang,supra)。

[1076] 结果表明,给药单一剂量的dsRNA剂抑制含重复序列的C9orf 72 mRNA的产生,但对于总C9orf72 mRNA水平的影响很小,而对于含外显子1b的mRNA水平无效。

[1077] 为了测定dsRNA剂降低含C9orf72正义链及反义链的团簇两者的数目和/或形成,采用Jiang,supra中描述的FISH方法于从给药来自上文的感兴趣的双链体的动物获得的样本中进行测定。所使用的探针包括针对正义及反义RNA六核苷酸重复序列的那些(Exiqon,Inc.)。全部杂交步骤都在无RNase条件下执行。将15微米的脑及脊髓OCT冷冻切片透化并将这些切片封闭。然后,这些切片与变性探针杂交。杂交之后,洗涤载玻片。将脂褐质的自发萦光淬灭,并且将细胞核用DAPI染色。由不知情的研究者执行对小鼠额叶皮质、海马齿状回、压后皮质及小脑分子层中的正义及反义RNA团簇的定量。通过共焦显微镜于100×放大下拍摄三至六张随机照片,并计数200至400个细胞。

[1078] 结果表明,给药单一剂量的dsRNA剂降低额叶皮质、海马齿状回、压后皮质及小脑

分子层中的正义及反义RNA团簇的水平。

[1079] 还评估了给药这些剂对于异常二肽重复蛋白质水平以及poly(GP)及poly(GA)负荷及尺寸的效应，如例如Jiang,supra中所描述，在给药上述感兴趣的双链体的动物中进行。

[1080] 使用免疫组织化学鉴定及评估小鼠半脑及脊髓中异常二肽重复蛋白质水平。简而言之，从福尔马林固定石蜡包埋块中切割8至10微米厚至小鼠半脑矢状切片或脊髓的冠状切片，并封固于载玻片上。干燥后，将载玻片于二甲苯和酒精中去石蜡化与再水化，然后洗涤。然后，载玻片用蒸汽处理并封闭。用可商购的抗poly(GP)、poly(GA)、poly(GR)、poly(PA)、poly(PR)、GFAP、IBA-1、CD3、F4/80及CD45R/B220抗体染色过夜之后，施加HRP缀合的二级抗体，并使用底物显影过氧化酶活性。切片用Harris氏经修饰的苏木精进一步染色并盖上盖玻片。

[1081] 为了定量poly(GP)及poly(GA)包合物负荷及尺寸，将已经针对poly(GP)或poly(GA)染色的小鼠半脑切片于40×放大下扫描，以获得高分辨率的数字化图像。使用合适的软件，计数海马体或压后皮质中指定区域内包合物的数目。为了量测这些区域内包合物的尺寸，用显微镜于63×放大下拍摄照片。尽管每个包合物在给定场域中只分析一次，可拍摄该场域的多张图像，以确保仅对清晰的包合物进行分析。将图像打开并放大，并使用轮廓工具来追踪每个包合物以测定其面积(μm^2)。对于每只小鼠，以μm2为单位计算每个测试区域内的包合物的平均尺寸。

[1082] 数据是用来确定给药单一剂量的dsRNA剂是否降低异常二肽重复蛋白质水平，尤其是poly(GP)及poly(GA)包合物负荷及尺寸的水平。

[1083]

表 19. C9orf72 dsRNA 剂的未修饰正义和反义链序列（Ensembl 基因 ID: ENSG00000147894）

人类染色体 9 构建物 38 (hg38) 中的靶标配合物	正义序列 5' 至 3'	来自 Ensembl 基因 ID: ENSG00000147894 的正义来源名称	反义序列 5' 至 3'	来自 Ensembl 基因 ID: ENSG00000147894 的反义来源名称
(hg38) chr9:27573685-27573707:-	GUGUGUUUUUGUUUUCCCAU	ENSG00000147894_394-414_C21U_s	AUGGGAAAAACAAAAACACACAC	ENSG00000147894_392-414_G1A_as
(hg38) chr9:27573684-27573706:-	UGUGUUUUUGUUUUUCCCACU	ENSG00000147894_393-413_C21U_s	AGUGGGAAAAACAAAAACACACA	ENSG00000147894_391-413_G1A_as
(hg38) chr9:27573683-27573705:-	GUGUUUUUGUUUUUCCCACCU	ENSG00000147894_392-412_C21U_s	AGGUGGGAAAAACAAAAACACAC	ENSG00000147894_390-412_G1A_as
(hg38) chr9:27573681-27573703:-	GUUUUUGUUUUUCCCACCCUU	ENSG00000147894_390-410_C21U_s	AAGGGUGGGAAAAACAAAAACAC	ENSG00000147894_388-410_G1A_as
(hg38) chr9:27573664-27573686:-	CCUCUCUCCCCCACUACUUGCU	ENSG00000147894_373-393_s	AGCAAGUAGUGGGGAGAGAGGGU	ENSG00000147894_371-393_as
(hg38) chr9:27573660-27573682:-	UCUCCCCACUACUUGCUCUCA	ENSG00000147894_369-389_s	UGAGAGCAAGUAGUGGGGAGAGA	ENSG00000147894_367-389_as
(hg38) chr9:27573659-27573681:-	CUCCCCACUACUUGCUCUCAU	ENSG00000147894_368-388_C21U_s	AUGAGAGCAAGUAGUGGGGAGAG	ENSG00000147894_366-388_G1A_as

[1084]

人类染色体 9 构建物 38 (hg38) 中的靶标配合物	正义序列 5' 至 3'	来自 Ensembl 基因 ID: ENSG00000147894 的正义来源名称	反义序列 5' 至 3'	来自 Ensembl 基因 ID: ENSG00000147894 的反义来源名称
(hg38) chr9:27573658-27573680:-	UCCCCACUACUUGCUCUCACA	ENSG00000147894_367-387_s	UGUGAGAGCAAGUAGUGGGAGA	ENSG00000147894_365-387_as
(hg38) chr9:27573657-27573679:-	CCCCACUACUUGCUCUCACAU	ENSG00000147894_366-386_G21U_s	AUGUGAGAGCAAGUAGUGGGAG	ENSG00000147894_364-386_C1A_as
(hg38) chr9:27573656-27573678:-	CCCACUACUUGCUCUCACAGU	ENSG00000147894_365-385_s	ACUGUGAGAGCAAGUAGUGGGGA	ENSG00000147894_363-385_as
(hg38) chr9:27573655-27573677:-	CCACUACUUGCUCUCACAGUA	ENSG00000147894_364-384_s	UACUGUGAGAGCAAGUAGUGGGG	ENSG00000147894_362-384_as
(hg38) chr9:27573654-27573676:-	CACUACUUGCUCUCACAGUAU	ENSG00000147894_363-383_C21U_s	AUACUGUGAGAGCAAGUAGUGGG	ENSG00000147894_361-383_G1A_as
(hg38) chr9:27573653-27573675:-	ACUACUUGCUCUCACAGUACU	ENSG00000147894_362-382_s	AGUACUGUGAGAGCAAGUAGUGG	ENSG00000147894_360-382_as
(hg38) chr9:27573652-27573674:-	CUACUUGCUCUCACAGUACUU	ENSG00000147894_361-381_C21U_s	AAGUACUGUGAGAGCAAGUAGUG	ENSG00000147894_359-381_G1A_as
(hg38) chr9:27573651-27573673:-	UACUUGCUCUCACAGUACUCU	ENSG00000147894_360-380_G21U_s	AGAGUACUGUGAGAGCAAGUAGU	ENSG00000147894_358-380_C1A_as

[1085]

人类染色体 9 构建物 38 (hg38) 中的靶标配合物	正义序列 5' 至 3'	来自 Ensembl 基因 ID: ENSG00000147894 的正义来源名称	反义序列 5' 至 3'	来自 Ensembl 基因 ID: ENSG00000147894 的反义来源名称
73673:-				
(hg38) chr9:27573650-27573672:-	ACUUGCUCUCACAGUACUCGU	ENSG00000147894_359-379_C21U_s	ACGAGUACUGUGAGAGCAAGUAG	ENSG00000147894_357-379_G1A_as
(hg38) chr9:27573649-27573671:-	CUUGCUCUCACAGUACUCGCU	ENSG00000147894_358-378_s	AGCGAGUACUGUGAGAGCAAGUA	ENSG00000147894_356-378_as
(hg38) chr9:27573648-27573670:-	UUGCUCUCACAGUACUCGCUU	ENSG00000147894_357-377_G21U_s	AAGCGAGUACUGUGAGAGCAAGU	ENSG00000147894_355-377_C1A_as
(hg38) chr9:27573647-27573669:-	UGCUCUCACAGUACUCGCUGA	ENSG00000147894_356-376_s	UCAGCGAGUACUGUGAGAGCAAG	ENSG00000147894_354-376_as
(hg38) chr9:27573646-27573668:-	GCUCUCACAGUACUCGCUGAU	ENSG00000147894_355-375_G21U_s	AUCAGCGAGUACUGUGAGAGCAA	ENSG00000147894_353-375_C1A_as
(hg38) chr9:27573645-27573667:-	CUCUCACAGUACUCGCUGAGU	ENSG00000147894_354-374_G21U_s	ACUCAGCGAGUACUGUGAGAGCA	ENSG00000147894_352-374_C1A_as
(hg38) chr9:27573643-27573665:-	CUCACAGUACUCGCUGAGGGU	ENSG00000147894_352-372_s	ACCCUCAGCGAGUACUGUGAGAG	ENSG00000147894_350-372_as
(hg38)	UCACAGUACUCGCUG	ENSG00000147894_351-371	AACCCUCAGCGAGUACU	ENSG00000147894_349-37

[1086]

人类染色体 9 构建物 38 (hg38) 中的靶标配合物	正义序列 5' 至 3'	来自 Ensembl 基因 ID: ENSG00000147894 的正义来源名称	反义序列 5' 至 3'	来自 Ensembl 基因 ID: ENSG00000147894 的反义来源名称
chr9:27573642-27573664:-	AGGGUU	1_G21U_s	GUGAGA	1_C1A_as
(hg38) chr9:27573641-27573663:-	CACAGUACUCGCUGAGGGUGA	ENSG00000147894_350-370_s	UCACCCUCAGCGAGUACUGUGAG	ENSG00000147894_348-370_as
(hg38) chr9:27573640-27573662:-	ACAGUACUCGCUGAGGGUGAA	ENSG00000147894_349-369_s	UUCACCCUCAGCGAGUACUGUGA	ENSG00000147894_347-369_as
(hg38) chr9:27573639-27573661:-	CAGUACUCGCUGAGGGUGAAU	ENSG00000147894_348-368_C21U_s	AUUCACCCUCAGCGAGUACUGUG	ENSG00000147894_346-368_G1A_as
(hg38) chr9:27573638-27573660:-	AGUACUCGCUGAGGGUGAACA	ENSG00000147894_347-367_s	UGUUCACCCUCAGCGAGUACUGU	ENSG00000147894_345-367_as
(hg38) chr9:27573637-27573659:-	GUACUCGCUGAGGGUGAACAA	ENSG00000147894_346-366_s	UUGUUCACCCUCAGCGAGUACUG	ENSG00000147894_344-366_as
(hg38) chr9:27573636-27573658:-	UACUCGCUGAGGGUGAACAAU	ENSG00000147894_345-365_G21U_s	AUUGUUCACCCUCAGCGAGUACU	ENSG00000147894_343-365_C1A_as
(hg38) chr9:27573635-27573657:-	ACUCGCUGAGGGUGAACAAGA	ENSG00000147894_344-364_s	UCUUGUUCACCCUCAGCGAGUAC	ENSG00000147894_342-364_as

[1087]

人类染色体 9 构建物 38 (hg38) 中的靶标配合物	正义序列 5' 至 3'	来自 Ensembl 基因 ID: ENSG00000147894 的正义来源名称	反义序列 5' 至 3'	来自 Ensembl 基因 ID: ENSG00000147894 的反义来源名称
(hg38) chr9:27573634-27573656:-	CUCGCUGAGGGUGAACAAGAA	ENSG00000147894_343-363_s	UUCUUGUUCACCCUCAGCGAGUA	ENSG00000147894_341-363_as
(hg38) chr9:27573633-27573655:-	UCGCUGAGGGUGAACAAGAAA	ENSG00000147894_342-362_s	UUUCUUGUUCACCCUCAGCGAGU	ENSG00000147894_340-362_as
(hg38) chr9:27573632-27573654:-	CGCUGAGGGUGAACAAGAAAA	ENSG00000147894_341-361_s	UUUUCUUGUUCACCCUCAGCGAG	ENSG00000147894_339-361_as
(hg38) chr9:27573631-27573653:-	GCUGAGGGUGAACAAGAAAAU	ENSG00000147894_340-360_G21U_s	AUUUUCUUGUUCACCCUCAGCGA	ENSG00000147894_338-360_C1A_as
(hg38) chr9:27573630-27573652:-	CUGAGGGUGAACAAGAAAAGA	ENSG00000147894_339-359_s	UCUUUUCUUGUUCACCCUCAGCG	ENSG00000147894_337-359_as
(hg38) chr9:27573629-27573651:-	UGAGGGUGAACAAGAAAAGAU	ENSG00000147894_338-358_C21U_s	AUCUUUUCUUGUUCACCCUCAGC	ENSG00000147894_336-358_G1A_as
(hg38) chr9:27573628-27573650:-	GAGGGUGAACAAGAAAAGACU	ENSG00000147894_337-357_C21U_s	AGUCUUUUCUUGUUCACCCUCAG	ENSG00000147894_335-357_G1A_as
(hg38) chr9:27573627-27573	AGGGUGAACAAGAAAAGACCU	ENSG00000147894_336-356_s	AGGUCUUUUCUUGUUCACCCUCA	ENSG00000147894_334-356_as

[1088]

人类染色体 9 构建物 38 (hg38) 中的靶标配合物	正义序列 5' 至 3'	来自 Ensembl 基因 ID: ENSG00000147894 的正义来源名称	反义序列 5' 至 3'	来自 Ensembl 基因 ID: ENSG00000147894 的反义来源名称
73649:-				
(hg38) chr9:27573626-27573648:-	GGGUGAACAAGAAAAGACCUU	ENSG00000147894_335-355_G21U_s	AAGGUCUUUUCUUGUUCACCCUC	ENSG00000147894_333-355_C1A_as
(hg38) chr9:27573625-27573647:-	GGUGAACAAGAAAAGACCUGA	ENSG00000147894_334-354_s	UCAGGUCUUUUCUUGUUCACCCU	ENSG00000147894_332-354_as
(hg38) chr9:27573624-27573646:-	GUGAACAAGAAAAGACCUGAU	ENSG00000147894_333-353_s	AUCAGGUCUUUUCUUGUUCACCC	ENSG00000147894_331-353_as
(hg38) chr9:27573623-27573645:-	UGAACAAGAAAAGACCUGAUA	ENSG00000147894_332-352_s	UAUCAGGUCUUUUCUUGUUCACC	ENSG00000147894_330-352_as
(hg38) chr9:27573622-27573644:-	GAACAAGAAAAGACCUGAUAA	ENSG00000147894_331-351_s	UUAUCAGGUCUUUUCUUGUUCAC	ENSG00000147894_329-351_as
(hg38) chr9:27573621-27573643:-	AACAAGAAAAGACCUGAUAAA	ENSG00000147894_330-350_s	UUUAUCAGGUCUUUUCUUGUUCA	ENSG00000147894_328-350_as
(hg38) chr9:27573620-27573642:-	ACAAGAAAAGACCUGAUAAAU	ENSG00000147894_329-349_G21U_s	AUUUAUCAGGUCUUUUCUUGUUC	ENSG00000147894_327-349_C1A_as
(hg38)	AAGAAAAGACCUGAU	ENSG00000147894_327-34	AUCUUUAUCAGGUCUU	ENSG00000147894_325-34

[1089]

人类染色体 9 构建物 38 (hg38) 中的靶标配合物	正义序列 5' 至 3'	来自 Ensembl 基因 ID: ENSG00000147894 的正义来源名称	反义序列 5' 至 3'	来自 Ensembl 基因 ID: ENSG00000147894 的反义来源名称
chr9:27573618-27573640:-	AAAGAU	7_s	UUCUUGU	7_as
(hg38) chr9:27573617-27573639:-	AGAAAAGACCUGAUAAAGAUU	ENSG00000147894_326-346_s	AAUCUUUAUCAGGUCUUUCUUG	ENSG00000147894_324-346_as
(hg38) chr9:27573616-27573638:-	GAAAAGACCUGAUAAAGAUUA	ENSG00000147894_325-345_s	UAAUCUUUAUCAGGUCUUUCUU	ENSG00000147894_323-345_as
(hg38) chr9:27573615-27573637:-	AAAAGACCUGAUAAAGAUUAA	ENSG00000147894_324-344_s	UUAAUCUUUAUCAGGUCUUUCU	ENSG00000147894_322-344_as
(hg38) chr9:27573614-27573636:-	AAAGACCUGAUAAAGAUUAAU	ENSG00000147894_323-343_C21U_s	AUUAAUCUUUAUCAGGUCUUUC	ENSG00000147894_321-343_G1A_as
(hg38) chr9:27573613-27573635:-	AAGACCUGAUAAAGAUUAACU	ENSG00000147894_322-342_C21U_s	AGUUAAUCUUUAUCAGGUCUUU	ENSG00000147894_320-342_G1A_as
(hg38) chr9:27573612-27573634:-	AGACCUGAUAAAGAUUAACCA	ENSG00000147894_321-341_1_s	UGGUUAAUCUUUAUCAGGUCUU	ENSG00000147894_319-341_1_as
(hg38) chr9:27573611-27573633:-	GACCUGAUAAAGAUUAACCAU	ENSG00000147894_320-340_G21U_s	AUGGUUAAUCUUUAUCAGGUCU	ENSG00000147894_318-340_C1A_as

[1090]

人类染色体 9 构建物 38 (hg38) 中的靶标配合物	正义序列 5' 至 3'	来自 Ensembl 基因 ID: ENSG00000147894 的正义来源名称	反义序列 5' 至 3'	来自 Ensembl 基因 ID: ENSG00000147894 的反义来源名称
(hg38) chr9:27573610-27573632:-	ACCUGAUAAAGAUUAACCAGA	ENSG00000147894_319-339_s	UCUGGUUAAUCUUUAUCAGGUCU	ENSG00000147894_317-339_as
(hg38) chr9:27573609-27573631:-	CCUGAUAAAGAUUAACCAGAA	ENSG00000147894_318-338_s	UUCUGGUUAAUCUUUAUCAGGUC	ENSG00000147894_316-338_as
(hg38) chr9:27573608-27573630:-	CUGAUAAAGAUUAACCAGAAU	ENSG00000147894_317-337_G21U_s	AUUCUGGUUAAUCUUUAUCAGGU	ENSG00000147894_315-337_C1A_as
(hg38) chr9:27573607-27573629:-	UGAUAAAGAUUAACCAGAAGA	ENSG00000147894_316-336_s	UCUUCUGGUUAAUCUUUAUCAGG	ENSG00000147894_314-336_as
(hg38) chr9:27573606-27573628:-	GAUAAAGAUUAACCAGAAGAA	ENSG00000147894_315-335_s	UUCUUCUGGUUAAUCUUUAUCAG	ENSG00000147894_313-335_as
(hg38) chr9:27573605-27573627:-	AUAAAGAUUAACCAGAAGAAA	ENSG00000147894_314-334_s	UUUCUUCUGGUUAAUCUUUAUCA	ENSG00000147894_312-334_as
(hg38) chr9:27573603-27573625:-	AAAGAUUAACCAGAAGAAAAU	ENSG00000147894_312-332_C21U_s	AUUUUCUUCUGGUUAAUCUUUAU	ENSG00000147894_310-332_G1A_as
(hg38) chr9:27573602-27573624:-	AAGAUUAACCAGAAGAAAACA	ENSG00000147894_311-331_s	UGUUUUCUUCUGGUUAAUCUUUA	ENSG00000147894_309-331_as

[1091]

人类染色体 9 构建物 38 (hg38) 中的靶标配合物	正义序列 5' 至 3'	来自 Ensembl 基因 ID: ENSG00000147894 的正义来源名称	反义序列 5' 至 3'	来自 Ensembl 基因 ID: ENSG00000147894 的反义来源名称
73624:-				
(hg38) chr9:27573599-27573621:-	AUUAACCAGAAGAAAACAAGU	ENSG00000147894_308-328_G21U_s	ACUUGUUUUCUUCUGGUUAAUCU	ENSG00000147894_306-328_C1A_as
(hg38) chr9:27573598-27573620:-	UUAACCAGAAGAAAACAAGGA	ENSG00000147894_307-327_s	UCCUUGUUUUCUUCUGGUUAAUC	ENSG00000147894_305-327_as
(hg38) chr9:27573597-27573619:-	UAACCAGAAGAAAACAAGGAU	ENSG00000147894_306-326_G21U_s	AUCCUUGUUUUCUUCUGGUUAAU	ENSG00000147894_304-326_C1A_as
(hg38) chr9:27573596-27573618:-	AACCAGAAGAAAACAAGGAGU	ENSG00000147894_305-325_G21U_s	ACUCCUUGUUUUCUUCUGGUUAA	ENSG00000147894_303-325_C1A_as
(hg38) chr9:27573595-27573617:-	ACCAGAAGAAAACAAGGAGGU	ENSG00000147894_304-324_G21U_s	ACCUCCUUGUUUUCUUCUGGUUA	ENSG00000147894_302-324_C1A_as
(hg38) chr9:27573577-27573599:-	GGGAAACAACCGCAGCCUGUA	ENSG00000147894_286-306_s	UACAGGCUGCGGUUGUUCCCUC	ENSG00000147894_284-306_as
(hg38) chr9:27573576-27573598:-	GGAAACAACCGCAGCCUGUAU	ENSG00000147894_285-305_G21U_s	AUACAGGCUGCGGUUGUUCCCU	ENSG00000147894_283-305_C1A_as
(hg38)	AAACAACCGCAGCCU	ENSG00000147894_283-303	UGCUACAGGCUGCGGUU	ENSG00000147894_281-303

[1092]

人类染色体 9 构建物 38 (hg38) 中的靶标配合物	正义序列 5' 至 3'	来自 Ensembl 基因 ID: ENSG00000147894 的正义来源名称	反义序列 5' 至 3'	来自 Ensembl 基因 ID: ENSG00000147894 的反义来源名称
chr9:27573574-27573596:-	GUAGCA	3_s	GUUUCC	3_as
(hg38) chr9:27573573-27573595:-	AACAACCGCAGCCUGUAGCAA	ENSG00000147894_282-302_s	UUGCUACAGGCUGCGGUUGUUUC	ENSG00000147894_280-302_as
(hg38) chr9:27573572-27573594:-	ACAACCGCAGCCUGUAGCAAU	ENSG00000147894_281-301_G21U_s	AUUGCUACAGGCUGCGGUUGUUU	ENSG00000147894_279-301_C1A_as
(hg38) chr9:27573571-27573593:-	CAACCGCAGCCUGUAGCAAGU	ENSG00000147894_280-300_C21U_s	ACUUGCUACAGGCUGCGGUUGUU	ENSG00000147894_278-300_G1A_as
(hg38) chr9:27573570-27573592:-	AACCGCAGCCUGUAGCAAGCU	ENSG00000147894_279-299_s	AGCUUGCUACAGGCUGCGGUUGU	ENSG00000147894_277-299_as
(hg38) chr9:27573568-27573590:-	CCGCAGCCUGUAGCAAGCUCU	ENSG00000147894_277-297_s	AGAGCUUGCUACAGGCUGCGGUU	ENSG00000147894_275-297_as
(hg38) chr9:27573564-27573586:-	AGCCUGUAGCAAGCUCUGGAA	ENSG00000147894_273-293_s	UUCCAGAGCUUGCUACAGGCUGC	ENSG00000147894_271-293_as
(hg38) chr9:27573563-27573585:-	GCCUGUAGCAAGCUCUGGAAU	ENSG00000147894_272-292_C21U_s	AUUCCAGAGCUUGCUACAGGCUG	ENSG00000147894_270-292_G1A_as

[1093]

人类染色体 9 构建物 38 (hg38) 中的靶标配合物	正义序列 5’ 至 3’	来自 Ensembl 基因 ID: ENSG00000147894 的正义来源名称	反义序列 5’ 至 3’	来自 Ensembl 基因 ID: ENSG00000147894 的反义来源名称
(hg38) chr9:27573562-27573584:-	CCUGUAGCAAGCUCUGGAACU	ENSG00000147894_271-291_s	AGUUCCAGAGCUUGCUACAGGCU	ENSG00000147894_269-291_as
(hg38) chr9:27573561-27573583:-	CUGUAGCAAGCUCUGGAACUU	ENSG00000147894_270-290_C21U_s	AAGUUCCAGAGCUUGCUACAGGC	ENSG00000147894_268-290_G1A_as
(hg38) chr9:27573560-27573582:-	UGUAGCAAGCUCUGGAACUCA	ENSG00000147894_269-289_s	UGAGUUCCAGAGCUUGCUACAGG	ENSG00000147894_267-289_as
(hg38) chr9:27573559-27573581:-	GUAGCAAGCUCUGGAACUCAU	ENSG00000147894_268-288_G21U_s	AUGAGUUCCAGAGCUUGCUACAG	ENSG00000147894_266-288_C1A_as
(hg38) chr9:27573558-27573580:-	UAGCAAGCUCUGGAACUCAGU	ENSG00000147894_267-287_G21U_s	ACUGAGUUCCAGAGCUUGCUACA	ENSG00000147894_265-287_C1A_as
(hg38) chr9:27573557-27573579:-	AGCAAGCUCUGGAACUCAGGA	ENSG00000147894_266-286_s	UCCUGAGUUCCAGAGCUUGCUAC	ENSG00000147894_264-286_as
(hg38) chr9:27573556-27573578:-	GCAAGCUCUGGAACUCAGGAU	ENSG00000147894_265-285_G21U_s	AUCCUGAGUUCCAGAGCUUGCUA	ENSG00000147894_263-285_C1A_as
(hg38) chr9:27573555-27573577:-	CAAGCUCUGGAACUCAGGAGU	ENSG00000147894_264-284_s	ACUCCUGAGUUCCAGAGCUUGCU	ENSG00000147894_262-284_as

[1094]

人类染色体 9 构建物 38 (hg38) 中的靶标配合物	正义序列 5' 至 3'	来自 Ensembl 基因 ID: ENSG00000147894 的正义来源名称	反义序列 5' 至 3'	来自 Ensembl 基因 ID: ENSG00000147894 的反义来源名称
73577:-				
(hg38) chr9:27573554-27573576:-	AAGCUCUGGAACUCAGGAGUU	ENSG00000147894_263-283_C21U_s	AACUCCUGAGUUCCAGAGCUUGC	ENSG00000147894_261-283_G1A_as
(hg38) chr9:27573553-27573575:-	AGCUCUGGAACUCAGGAGUCU	ENSG00000147894_262-282_G21U_s	AGACUCCUGAGUUCCAGAGCUUG	ENSG00000147894_260-282_C1A_as
(hg38) chr9:27573552-27573574:-	GCUCUGGAACUCAGGAGUCGU	ENSG00000147894_261-281_C21U_s	ACGACUCCUGAGUUCCAGAGCUU	ENSG00000147894_259-281_G1A_as
(hg38) chr9:27573542-27573564:-	UCAGGAGUCGCGCGCUAGGGU	ENSG00000147894_251-271_G21U_s	ACCCUAGCGCGCGACUCCUGAGU	ENSG00000147894_249-271_C1A_as
(hg38) chr9:27573541-27573563:-	CAGGAGUCGCGCGCUAGGGGU	ENSG00000147894_250-270_C21U_s	ACCCCUAGCGCGCGACUCCUGAG	ENSG00000147894_248-270_G1A_as
(hg38) chr9:27573540-27573562:-	AGGAGUCGCGCGCUAGGGCU	ENSG00000147894_249-269_C21U_s	AGCCCCUAGCGCGCGACUCCUGA	ENSG00000147894_247-269_G1A_as
(hg38) chr9:27573539-27573561:-	GGAGUCGCGCGCUAGGGCCU	ENSG00000147894_248-268_G21U_s	AGGCCCCUAGCGCGCGACUCCUG	ENSG00000147894_246-268_C1A_as
(hg38)	GAGUCGCGCGCUAGG	ENSG00000147894_247-267	ACGGCCCCUAGCGCGCG	ENSG00000147894_245-267

[1095]

人类染色体 9 构建物 38 (hg38) 中的靶标配合物	正义序列 5’至 3’	来自 Ensembl 基因 ID: ENSG00000147894 的正义来源名称	反义序列 5’至 3’	来自 Ensembl 基因 ID: ENSG00000147894 的反义来源名称
chr9:27573538-27573560:-	GGCCGU	7_G21U_s	ACUCCU	7_C1A_as
(hg38) chr9:27573537-27573559:-	AGUCGCGCGCUAGGGGCCGGU	ENSG00000147894_246-266_G21U_s	ACCGGCCCCUAGCGCGCGACUCC	ENSG00000147894_244-266_C1A_as

[1096]

表 20. C9orf72 dsRNA 剂的修饰正义和反义链序列（Ensembl 基因 ID: ENSG00000147894）

人类染色体 9 构建物 38 (hg38) 中的靶标配合物	靶标 mRNA 序列	反义序列 5' 至 3'	正义序列 5' 至 3'
(hg38) chr9:27573685-27573707:-AC	GUGUGUGUUUUUGUUUUUCCC	asUfsgggAfaAfAfacaaAfaAfAfcacacsasc	gsusguguUfuUfUfGfuuuuucccauL96
(hg38) chr9:27573684-27573706:-CC	UGUGUGUUUUUGUUUUUCCCA	asGfsuggGfaAfAfaacaAfaAfAfacacascsa	usgsuguuUfuUfGfUfuuuucccacuL96
(hg38) chr9:27573683-27573705:-CC	GUGUGUUUUUGUUUUUCCCAC	asGfsgugGfgAfAfaaacAfaAfAfaacacsasc	gsusguuuUfuGfUfUfuuucccaccuL96
(hg38) chr9:27573681-27573703:-UC	GUGUUUUUGUUUUUCCCACCC	asAfsgggUfgGfGfaaaaAfcAfAfaaaacsasc	gsusuuuuGfuUfUfUfucccaccccuuL96
(hg38) chr9:27573664-27573686:-U	ACCCUCUCUCCCCCACUACUUGC	asGfscaaGfuAfGfugggGfaGfAfgaggsgsu	cscsucucUfcCfCfCfacuacuugcuL96
(hg38) chr9:27573660-27573682:-CA	UCUCUCCCCACUACUUGCUCU	usGfsagaGfcAfAfguagUfgGfggagasgsa	uscsucccCfaCfUfAfcuugcucucaL96
(hg38) chr9:27573659-27573681:-AC	CUCUCCCCACUACUUGCUCUC	asUfsgagAfgCfAfaguaGfuGfgggagsasg	csuscccccAfcUfAfCfuugcucucauL96
(hg38) chr9:27573658-27573680:-CA	UCUCCCCACUACUUGCUCUCA	usGfsugaGfaGfCfaaguAfgUfggggasgsa	uscscccaCfuAfCfUfugcucucacaL96
(hg38) chr9:27573657-27573679:-AG	CUCCCCACUACUUGCUCUCAC	asUfsgugAfgAfGfcaagUfaGfuggggsasg	cscsccacUfaCfUfUfgcucucacauL96
(hg38) chr9:27573656-27573678:-GU	UCCCCACUACUUGCUCUCACA	asCfsuguGfaGfAfgcaaGfuAfgugggsgsa	cscscacuAfcUfUfGfcucucacaguL96
(hg38) chr9:27573655-27573677:-UA	CCCCACUACUUGCUCUCACAG	usAfscugUfgAfGfagcaAfgUfaguggsgsg	cscsacuaCfuUfGfCfucucacaguaL96

[1097]

人类染色体 9 构建物 38 (hg38) 中的靶标配合物	靶标 mRNA 序列	反义序列 5' 至 3'	正义序列 5' 至 3'
(hg38) chr9:27573654-27573676:-AC	CCCACUACUUGCUCUCACAGU	asUfsacuGfuGfAfgagcAfaGfuagugsgsg	csascuacUfuGfCfUfcucacaguauL96
(hg38) chr9:27573653-27573675:-CU	CCACUACUUGCUCUCACAGUA	asGfsuacUfgUfGfagagCfaAfguagusgsg	ascsuacuUfgCfUfCfucacaguacuL96
(hg38) chr9:27573652-27573674:-UC	CACUACUUGCUCUCACAGUAC	asAfsguaCfuGfUfgagaGfcAfaguagsusg	csusacuuGfcUfCfUfcacaguacuuL96
(hg38) chr9:27573651-27573673:-CG	ACUACUUGCUCUCACAGUACU	asGfsaguAfcUfGfugagAfgCfaaguasgsu	usascuugCfuCfUfCfacaguacucuL96
(hg38) chr9:27573650-27573672:-GC	CUACUUGCUCUCACAGUACUC	asCfsgagUfaCfUfgugaGfaGfcaagusasg	ascsuugcUfcUfCfUfcaguacucguL96
(hg38) chr9:27573649-27573671:-CU	UACUUGCUCUCACAGUACUCG	asGfscgaGfuAfCfugugAfgAfgcaagsusa	csusugcuCfuCfAfCfaguacucgcuL96
(hg38) chr9:27573648-27573670:-UG	ACUUGCUCUCACAGUACUCGC	asAfsgcgAfgUfAfcuguGfaGfagcaasgsu	ususgcucUfcAfCfAfguacucgcunL96
(hg38) chr9:27573647-27573669:-GA	CUUGCUCUCACAGUACUCGCU	usCfsagcGfaGfUfacugUfgAfgagcasasg	usgscucuCfaCfAfGfuacucgcugaL96
(hg38) chr9:27573646-27573668:-AG	UUGCUCUCACAGUACUCGCUG	asUfscagCfgAfGfuacuGfuGfagagcsasa	gscsucucAfcAfGfUfacucgcugauL96
(hg38) chr9:27573645-27573667:-GG	UGCUCUCACAGUACUCGCUGA	asCfsucaGfcGfAfguacUfgUfgagagscsa	csuscucaCfaGfUfAfcucgcugaguL96
(hg38) chr9:27573643-27573665:-GU	CUCUCACAGUACUCGCUGAGG	asCfsccuCfaGfCfgaguAfcUfgugagsasg	csuscacaGfuAfCfUfcgcugaggguL96
(hg38) chr9:27573642-27573664:-UG	UCUCACAGUACUCGCUGAGGG	asAfscccUfcAfGfcgagUfaCfugugasgsa	uscsacagUfaCfUfCfgcugaggguuL96

[1098]

人类染色体 9 构建物 38 (hg38) 中的靶标配合物	靶标 mRNA 序列	反义序列 5' 至 3'	正义序列 5' 至 3'
(hg38) chr9:27573641-27573663:-GA	CUCACAGUACUCGCUGAGGGUGA	usCfsaccCfuCfAfgcgaGfuAfcugugsasg	csascaguAfcUfCfGfcugagggugaL96
(hg38) chr9:27573640-27573662:-AA	UCACAGUACUCGCUGAGGGUG	usUfscacCfcUfCfagcgAfgUfacugusgsa	ascsaguaCfuCfGfCfugagggugaaL96
(hg38) chr9:27573639-27573661:-AC	CACAGUACUCGCUGAGGGUGA	asUfsucaCfcCfUfcagcGfaGfuacugsusg	csasguacUfcGfCfUfgagggugaauL96
(hg38) chr9:27573638-27573660:-CA	ACAGUACUCGCUGAGGGUGAA	usGfsuucAfcCfCfucagCfgAfguacusgsu	asgsuacuCfgCfUfGfagggugaacaL96
(hg38) chr9:27573637-27573659:-AA	CAGUACUCGCUGAGGGUGAAC	usUfsguuCfaCfCfcucaGfcGfaguacsusg	gsusacucGfcUfGfAfgggugaacaaL96
(hg38) chr9:27573636-27573658:-AG	AGUACUCGCUGAGGGUGAACA	asUfsuguUfcAfCfccucAfgCfgaguascsu	usascucgCfuGfAfGfgggugaacaauL96
(hg38) chr9:27573635-27573657:-GA	GUACUCGCUGAGGGUGAACAA	usCfsuugUfuCfAfcccuCfaGfcgagusasc	ascsucgcUfgAfGfGfgugaacaagaL96
(hg38) chr9:27573634-27573656:-AA	UACUCGCUGAGGGUGAACAAG	usUfscuuGfuUfCfacccUfcAfgcgagsusa	csuscgcuGfaGfGfGfugaacaagaaL96
(hg38) chr9:27573633-27573655:-AA	ACUCGCUGAGGGUGAACAAGA	usUfsucuUfgUfUfcaccCfuCfagcgasgsu	uscsgcugAfgGfGfUfgaacaagaaaL96
(hg38) chr9:27573632-27573654:-AA	CUCGCUGAGGGUGAACAAGAA	usUfsuucUfuGfUfucacCfcUfcagcgsasg	csgscugaGfgGfUfGfaacaagaaaaL96
(hg38) chr9:27573631-27573653:-AG	UCGCUGAGGGUGAACAAGAAA	asUfsuuuCfuUfGfuucaCfcCfucagcsgsa	gscsugagGfgUfGfAfacaagaaaauL96
(hg38) chr9:27573630-27573652:-GA	CGCUGAGGGUGAACAAGAAAA	usCfsuuuUfcUfUfguucAfcCfcucagscsg	csusgaggGfuGfAfAfcaagaaaagaL96

[1099]

人类染色体 9 构建物 38 (hg38) 中的靶标配合物	靶标 mRNA 序列	反义序列 5' 至 3'	正义序列 5' 至 3'
(hg38) chr9:27573629-27573651:-AC	GCUGAGGGUGAACAAGAAAAG	asUfscuuUfuCfUfuguuCfaCfccucas gsc	usgsagggUfgAfAfCfaagaaaagauL 96
(hg38) chr9:27573628-27573650:-CC	CUGAGGGUGAACAAGAAAAGA	asGfsucuUfuUfCfuugUfcAfcccucs asg	gsasggguGfaAfCfAfagaaaagacuL 96
(hg38) chr9:27573627-27573649:-CU	UGAGGGUGAACAAGAAAAGAC	asGfsgucUfuUfUfcuugUfuCfacccus csa	asgsggugAfaCfAfAfgaaaagaccuL 96
(hg38) chr9:27573626-27573648:-UG	GAGGGUGAACAAGAAAAGACC	asAfsgguCfuUfUfucuuGfuUfcacccs usc	gsgsgugaAfcAfAfGfaaaagaccuuL 96
(hg38) chr9:27573625-27573647:-GA	AGGGUGAACAAGAAAAGACCU	usCfsaggUfcUfUfuucuUfgUfucaccs csu	gsgsugaaCfaAfGfAfaaagaccugaL 96
(hg38) chr9:27573624-27573646:-AU	GGGUGAACAAGAAAAGACCUG	asUfscagGfuCfUfuuucUfuGfuucacs csc	gsusgaacAfaGfAfAfaagaccugauL 96
(hg38) chr9:27573623-27573645:-UA	GGUGAACAAGAAAAGACCUGA	usAfsucaGfgUfCfuuuuCfuUfguucas csc	usgsaacaAfgAfAfAfagaccugauaL 96
(hg38) chr9:27573622-27573644:-AA	GUGAACAAGAAAAGACCUGAU	usUfsaucAfgGfUfcuuuUfcUfuguucs asc	gsasacaaGfaAfAfAfgaccugauaaL 96
(hg38) chr9:27573621-27573643:-AA	UGAACAAGAAAAGACCUGAUA	usUfsuauCfaGfGfucuuUfuCfuuguus csa	asascaagAfaAfAfGfaccugauaaaL 96
(hg38) chr9:27573620-27573642:-AG	GAACAAGAAAAGACCUGAUAA	asUfsuuaUfcAfGfgucuUfuUfcuugus usc	ascsaagaAfaAfGfAfccugauaaauL 96
(hg38) chr9:27573618-27573640:-AU	ACAAGAAAAGACCUGAUAAAG	asUfscuuUfaUfCfagguCfuUfuucuus gsu	asasgaaaAfgAfCfCfugauaaagauL 96
(hg38) chr9:27573617-27573639:-UU	CAAGAAAAGACCUGAUAAAGA	asAfsucuUfuAfUfcaggUfcUfuuucus usg	asgsaaaaGfaCfCfUfgauaaagauuL 96

[1100]

人类染色体 9 构建物 38 (hg38) 中的靶标配合物	靶标 mRNA 序列	反义序列 5' 至 3'	正义序列 5' 至 3'
(hg38) chr9:27573616-27573638:-UA	AAGAAAAGACCUGAUAAAGAU	usAfsaucUfuUfAfucagGfuCfuuuucs usu	gsasaaagAfcCfUfGfauaaaganuaL 96
(hg38) chr9:27573615-27573637:-AA	AGAAAAGACCUGAUAAAGAUU	usUfsaauCfuUfUfaucaGfgUfcuuuus csu	asasaagaCfcUfGfAfuaaagauuaaL 96
(hg38) chr9:27573614-27573636:-AC	GAAAAGACCUGAUAAAGAUUA	asUfsuaaUfcUfUfuaucAfgGfucuuus usc	asasagacCfuGfAfUfaaagauuaauL 96
(hg38) chr9:27573613-27573635:-CC	AAAAGACCUGAUAAAGAUUAA	asGfsuuaAfuCfUfuuauCfaGfgucuus usu	asasgaccUfgAfUfAfaagauuaacuL 96
(hg38) chr9:27573612-27573634:-CA	AAAGACCUGAUAAAGAUUAAC	usGfsguuAfaUfCfuuuaUfcAfggucus usu	asgsaccuGfaUfAfAfagauuaaccaL 96
(hg38) chr9:27573611-27573633:-AG	AAGACCUGAUAAAGAUUAACC	asUfsggnUfaAfUfcuuuAfuCfaggucs usu	gsasccugAfuAfAfAfgauuaaccauL 96
(hg38) chr9:27573610-27573632:-GA	AGACCUGAUAAAGAUUAACCA	usCfsuggUfuAfAfucuuUfaUfcaggus csu	ascscugaUfaAfAfGfauuaaccagaL 96
(hg38) chr9:27573609-27573631:-AA	GACCUGAUAAAGAUUAACCAG	usUfscugGfuUfAfaucuUfuAfucaggs usc	cscsugauAfaAfGfAfuuaaccagaaL 96
(hg38) chr9:27573608-27573630:-AG	ACCUGAUAAAGAUUAACCAGA	asUfsucuGfgUfUfaaucUfuUfaucags gsu	csusgauaAfaGfAfUfuaaccagaauL 96
(hg38) chr9:27573607-27573629:-GA	CCUGAUAAAGAUUAACCAGAA	usCfsuucUfgGfUfuaauCfuUfuaucas gsg	usgsauaaAfgAfUfUfaaccagaagaL 96
(hg38) chr9:27573606-27573628:-AA	CUGAUAAAGAUUAACCAGAAG	usUfscuuCfuGfGfuuaaUfcUfuuaucs asg	gsasuaaaGfaUfUfAfaccagaagaaL 96
(hg38) chr9:27573605-27573627:-AA	UGAUAAAGAUUAACCAGAAGA	usUfsucuUfcUfGfguuaAfuCfuuuaus csa	asusaaagAfuUfAfAfccagaagaaaL 96

[1101]

人类染色体 9 构建物 38 (hg38) 中的靶标配合物	靶标 mRNA 序列	反义序列 5' 至 3'	正义序列 5' 至 3'
(hg38) chr9:27573603-27573625:-	AUAAAGAUUAACCAGAAGAAA AC	asUfsuuuCfuUfCfugguUfaAfucuuus asu	asasganUfaAfCfCfagaagaaaauL 96
(hg38) chr9:27573602-27573624:-	UAAAGAUUAACCAGAAGAAAA CA	usGfsuuuUfcUfUfcuggUfuAfaucuus usa	asasgauuAfaCfCfAfgaagaaaacaL 96
(hg38) chr9:27573599-27573621:-	AGAUUAACCAGAAGAAAACAA GG	asCfsuugUfuUfUfcuucUfgGfuuaaus csu	asusuaacCfaGfAfAfgaaaacaaguL 96
(hg38) chr9:27573598-27573620:-	GAUUAACCAGAAGAAAACAAG GA	usCfscuuGfuUfUfUfucuuCfuGfguuaas usc	ususaaccAfgAfAfGfaaaacaaggaL 96
(hg38) chr9:27573597-27573619:-	AUUAACCAGAAGAAAACAAGG AG	asUfsccuUfgUfUfuucuUfcUfgguuas asu	usasaccaGfaAfGfAfaaacaaggauL 96
(hg38) chr9:27573596-27573618:-	UUAACCAGAAGAAAACAAGGA GG	asCfsuccUfuGfUfuuucUfuCfugguus asa	asasccagAfaGfAfAfaacaaggaguL 96
(hg38) chr9:27573595-27573617:-	UAACCAGAAGAAAACAAGGAG GG	asCfscucCfuUfGfuuuuCfuUfcuggus usa	ascscagaAfgAfAfAfacaaggaggnL 96
(hg38) chr9:27573577-27573599:-	GAGGGAAACAACCGCAGCCUG UA	usAfscagGfcUfGfcggnUfgUfuucccs usc	gsgsgaaaCfaAfCfCfgcagccuguaL 96
(hg38) chr9:27573576-27573598:-	AGGGAAACAACCGCAGCCUGU AG	asUfsacaGfgCfUfgcggUfuGfuuuccs csu	gsgsaaacAfaCfCfGfcagccuguauL 96
(hg38) chr9:27573574-27573596:-	GGAAACAACCGCAGCCUGUAG CA	usGfscuaCfaGfGfcugcGfgUfuguuus csc	asasacaaCfcGfCfAfgccuguagcaL 96
(hg38) chr9:27573573-27573595:-	GAAACAACCGCAGCCUGUAGC AA	usUfsgcuAfcAfGfgcugCfgGfuuguus usc	asascaacCfgCfAfGfccuguagcaaL 96
(hg38) chr9:27573572-27573594:-	AAACAACCGCAGCCUGUAGCA AG	asUfsugcUfaCfAfggcuGfcGfguugus usu	ascsaaccGfcAfGfCfcuguagcaauL 96

[1102]

人类染色体 9 构建物 38 (hg38) 中的靶标配合物	靶标 mRNA 序列	反义序列 5’ 至 3’	正义序列 5’ 至 3’
(hg38) chr9:27573571-27573593:-GC	AACAACCGCAGCCUGUAGCAA	asCfsuugCfuAfCfaggcUfgCfgguugsusu	csasaccgCfaGfCfCfuguagcaaguL96
(hg38) chr9:27573570-27573592:-CU	ACAACCGCAGCCUGUAGCAAG	asGfscuuGfcUfAfcaggCfuGfcgguusgsu	asasccgcAfgCfCfUfguagcaagcuL96
(hg38) chr9:27573568-27573590:-CU	AACCGCAGCCUGUAGCAAGCU	asGfsagcUfuGfCfuacaGfgCfugcggsusu	cscsgcagCfcUfGfUfagcaagcucuL96
(hg38) chr9:27573564-27573586:-AA	GCAGCCUGUAGCAAGCUCUGG	usUfsccaGfaGfCfuugcUfaCfaggcusgsc	asgsccugUfaGfCfAfagcucuggaaL96
(hg38) chr9:27573563-27573585:-AC	CAGCCUGUAGCAAGCUCUGGA	asUfsuccAfgAfGfcuugCfuAfcaggcsusg	gscscuguAfgCfAfAfgcucuggaauL96
(hg38) chr9:27573562-27573584:-CU	AGCCUGUAGCAAGCUCUGGAA	asGfsuucCfaGfAfgcuuGfcUfacaggscsu	cscsuguaGfcAfAfGfcucuggaacuL96
(hg38) chr9:27573561-27573583:-UC	GCCUGUAGCAAGCUCUGGAAC	asAfsguuCfcAfGfagcuUfgCfuacagsgsc	csusguagCfaAfGfCfucuggaacuuL96
(hg38) chr9:27573560-27573582:-CA	CCUGUAGCAAGCUCUGGAACU	usGfsagnUfcCfAfgagcUfuGfcuacasgsg	usgsuagcAfaGfCfUfcuggaacucaL96
(hg38) chr9:27573559-27573581:-AG	CUGUAGCAAGCUCUGGAACUC	asUfsgagUfuCfCfagagCfuUfgcuacsasg	gsusagcaAfgCfUfCfuggaacucauL96
(hg38) chr9:27573558-27573580:-GG	UGUAGCAAGCUCUGGAACUCA	asCfsugaGfuUfCfcagaGfcUfugcuascsa	usasgcaaGfcUfCfUfggaacucaguL96
(hg38) chr9:27573557-27573579:-GA	GUAGCAAGCUCUGGAACUCAG	usCfscugAfgUfUfccagAfgCfuugcusasc	asgscaagCfuCfUfGfgaacucaggaL96
(hg38) chr9:27573556-27573578:-AG	UAGCAAGCUCUGGAACUCAGG	asUfsccuGfaGfUfuccaGfaGfcuugcsusa	gscsaagcUfcUfGfUfggaacucaggauL96

[1103]

人类染色体 9 构建物 38 (hg38) 中的靶标配合物	靶标 mRNA 序列	反义序列 5' 至 3'	正义序列 5' 至 3'
(hg38) chr9:27573555-27573577:-GU	AGCAAGCUCUGGAACUCAGGA	asCfsuccUfgAfGfuuccAfgAfgcuugs csu	csasgcuCfuGfGfAfacucaggaguL 96
(hg38) chr9:27573554-27573576:-UC	GCAAGCUCUGGAACUCAGGAG	asAfscucCfuGfAfguucCfaGfagcuus gsc	asasgcucUfgGfAfAfcucaggaguuL 96
(hg38) chr9:27573553-27573575:-CG	CAAGCUCUGGAACUCAGGAGU	asGfsacuCfcUfGfaguuCfcAfgagcus usg	asgscucuGfgAfAfCfucaggagucuL 96
(hg38) chr9:27573552-27573574:-GC	AAGCUCUGGAACUCAGGAGUC	asCfsgacUfcCfUfgaguUfcCfagagcsu su	gscsucugGfaAfCfUfcaggagucguL 96
(hg38) chr9:27573542-27573564:-GG	ACUCAGGAGUCGCGCGCUAGG	asCfsccuAfgCfGfcgcgAfcUfccugasg su	uscsaggaGfuCfGfCfgcgcuaggguL 96
(hg38) chr9:27573541-27573563:-GC	CUCAGGAGUCGCGCGCUAGGG	asCfscccUfaGfCfgcgcGfaCfuccugsa sg	csasggagUfcGfCfGfcgcuagggguL 96
(hg38) chr9:27573540-27573562:-CC	UCAGGAGUCGCGCGCUAGGGG	asGfscccCfuAfGfcgcgCfgAfcuccusg sa	asgsgaguCfgCfGfCfgcuaggggcuL 96
(hg38) chr9:27573539-27573561:-CG	CAGGAGUCGCGCGCUAGGGGC	asGfsgccCfcUfAfgcgcGfcGfacuccsu sg	gsgsagucGfcGfCfGfcuaggggccuL 96
(hg38) chr9:27573538-27573560:-GG	AGGAGUCGCGCGCUAGGGGCC	asCfsggcCfcCfUfagcgCfgCfgacucsc su	gsasgucgCfgCfGfCfuaggggccguL 96
(hg38) chr9:27573537-27573559:-GG	GGAGUCGCGCGCUAGGGGCCG	asCfscggCfcCfCfuagcGfcGfgacusc sc	asgsucgcGfcGfCfUfaggggccgguL 96

[1104] 实施例7:靶向外显子1A、内含子1A及内含子1B的C9orf72 siRNA

[1105] 为进一步探索实施例4中所观察到的现象,携带具有300次重复的GGGGCC的人源化C9orf72基因的小鼠胚胎干细胞用靶向外显子1A、内含子1A(六核苷酸重复扩增的上游)及内含子1B(六核苷酸重复扩增的上游且外显子1B的下游)中C9orf72 RNA序列的siRNA处理。三种siRNA靶向外显子1A(siRNA E1A#1、E1A#2及E1A#3),四种siRNA靶向位于六核苷酸重复序列上游的内含子1A(siRNA I1A#2、1A#4、I1A#6及I1A#10),而两种siRNA靶向位于六核苷酸重复上游且位于外显子1B下游的内含子1B(siRNA I1B#1及I1B#6)。表21提供这些剂的正义链及反义链核苷酸序列。四种靶向内含子1A的siRNA所靶向的C9orf72区域以及实施例4中所使用的siRNA显示于图18A中。使用Gapdh siRNA作为阳性对照,使用关于Gapdh表达的

TAQMAN®定性PCR测定证实Gapdh表达的敲低(数据未显示)。使用模拟作为阴性对照。评估使用25nM siRNA转染后24小时、48小时及72小时的基因表达。然后,使用位于多个区域侧翼的引物及检测经修饰的C9orf72的那些区域的探针,执行TAQMAN®定性PCR测定。引物及探针显示于图18B中及表8中。这表明,靶向内含子1A中的内含子序列(GGGGCC重复序列扩增的上游)的siRNA促使含内含子1A的C9orf72 RNA的累积降低(测定F;参见,图19A、图20A和图21A,分别对应于24h、48h及72h时间点),而对于C9orf72成熟mRNA具有很低的效应(外显子2-3测定,ID Mm01216829m1,来自Thermo Fisher Scientific;参见,图19B、图20B和图21B,分别对应于24h、48h及72h时间点)。相比之下,靶向外显子1A或内含子1B(GGGGCC重复序列扩增的下游及外显子1B的下游)的siRNA并不降低含内含子1A的C9orf72 RNA的累积(参见,图19A、图20A和图21A,其分别对应于24h、48h及72h时间点)。该Gapdh siRNA对照显示,C9orf72 RNA的特异性敲低并非仅是细胞中常规RNA干扰活性的结果。

[1106] 将靶向C9orf72成熟RNA的siRNA的效应与靶向C9orf72内含子1A的siRNA的效应在携带具有300次重复的GGGGCC六核苷酸的人源化C9orf72基因的小鼠胚胎干细胞中进行比较。通过电穿孔将siRNA引入细胞内,并于电穿孔的后24小时及48小时评估基因表达。使用一种靶向C9orf72成熟RNA的siRNA(siRNA#1),并且使用一种靶向C9orf72内含子1A的siRNA(siRNA#12)。然后,使用位于多个区域侧翼的引物及检测经修饰的C9orf72的那些区域的探针,执行TAQMAN® 定性PCR测定。引物及探针显示于图18B及表8中。siRNA数据严密且再现了以下的先前结果:当靶向内含子1A时则选择性敲低含内含子1A的C9orf72 RNA(测定F)(参见,图22A及图23A中的siRNA#12),而对于C9orf72成熟RNA(测定D)仅具有很低的效应(参见,图22E及图23E中的siRNA#12)。靶向C9orf72成熟RNA的siRNA具有相反的效应。

[1107] 此外,靶向成熟mRNA本体的siRNA敲低C9orf72外显子1A-外显子2经剪接的mRNA(测定B),指示它们是共价连接的。相比之下,靶向C9orf72内含子1A(六核苷酸重复的上游)的siRNA并不显著影响外显子1A-外显子2经剪接的mRNA的表达。参见图22C及图23C。

[1108] 靶向内含子1A的siRNA也不敲低于外显子1B处启动的占优势的正常C9orf72 mRNA,如通过测定C结果所表明。参见图22D及图23D。测定C(外显子1B-外显子2剪接)表明,敲低含邻近六核苷酸重复的内含子1A的RNA(推定为致病性含重复的RNA)的siRNA保留正常mRNA,因此,并不显著降低从丰度最大形式的C9orf72 mRNA翻译的C9orf72蛋白质。

[1109] 测定E显示,靶向内含子1A的siRNA还敲低以下RNA,该RNA保留位于六核苷酸重复上游的内含子序列且连接至外显子1A。参见图22B及图23B。本类型的转录本将会是以下的:其中转录在外显子1A处启动,延伸穿过重复序列到达外显子1B,然后从外显子1B剪接至外显子2及mRNA的本体。此类RNA可用作由六核苷酸重复编码的二肽重复蛋白质合成的模板。我们研究的结果显示,靶向内含子1A的siRNA将会摧毁此类转录本并对DPR蛋白质合成的降低有所贡献,至少对于poly(Gly-Ala)是如此,我们使用这一siRNA治疗时观察到这一点。总之,靶向位于六核苷酸重复上游的内含子的siRNA保留成熟C9orf72mRNA。它们对于含内含子的C9orf72转录本(包括具有连接至相邻内含子的外显子1A的RNA),非常具有特异性。

[1110] 表21.

[1111]

靶标	名称	序列 5’ 至 3’
小鼠/人外显子 1A	siRNA_E1A#1	正义: 5' GGAAAGUGGGGUCUAGCAAUU 3' 反义: 5' UUGCUAGACCCCACUUUCCUU 3'
人外显子 1A	siRNA_E1A#2	正义: 5' CAAGAGCAGGUGUGGGUUUUU 3' 反义: 5' AAACCCACACCUGCUCUUGUU 3'
小鼠外显子 1A	siRNA_E1A#3	正义: 5' CCUUCUAGGUGGAAAGUGGUU 3' 反义: 5'-PCCACUUUCCACCUAGAAGGUU 3'
人内含子 1A	siRNA_I1A#2	正义: 5' CAAGAAAAGACCUGAUAAAUU 3' 正义: 5' UUUAUCAGGUCUUUUCUUGUU 3'
	siRNA_I1A#4	正义: 5' GAAAAGACCUGAUAAAGAUUU 3' 反义: 5' AUCUUUAUCAGGUCUUUUCUU 3'
	siRNA_I1A#6	正义: 5' CGCUGAGGGUGAACAAGAAUU 3' 反义: 5' UUCUUGUUCACCCUCAGCGUU 3'
	siRNA_I1A#10	正义: 5' GAAGAAAACAAGGAGGGAAUU 3' 反义: 5' UUCCCUCCUUGUUUUCUUCUU 3'
人内含子 1B	siRNA_I1B#1	正义: 5' CCUCAGAGCUCGACGCAUUUU 3' 反义: 5' AAUGCGUCGAGCUCUGAGGUU 3'
	siRNA_I1B#6	正义: 5' UCUAGCGACUGGUGGAAUUUU 3' 反义: 5' AAUUCCACCAGUCGCUAGAUU 3'

非正式序列表

<210> 46
<211> 198
<212> DNA
<213> 智人 (Homo sapiens)

<400> 46
gggtctagca agagcaggtg tgggtttagg aggtgtgtgt ttttgttttt cccaccctct 60

ctccccacta cttgctctca cagtactcgc tgagggtgaa caagaaaaga cctgataaag 120

attaaccaga agaaaacaag gagggaaaca accgcagcct gtagcaagct ctggaactca 180

ggagtcgcgc gctatgcg 198

<210> 47
<211> 118
<212> DNA
<213> 智人 (Homo sapiens)

<400> 47
gcgatcgcgg ggcgtggtcg gggcgggccc gggggcgggc ccggggcggg gctgcggttg 60

cggtgcctgc gcccgcggcg gcggaggcgc aggcggtggc gagtgggtga gtgaggag 118

[1112]

<210> 111
<211> 6702
<212> RNA
<213> 人工序列

<220>
<223> 合成

<220>
<221> misc_feature
<222> (1)..(158)
<223> 外显子 1A

<220>
<221> misc_feature
<222> (321)..(338)
<223> 3x 六核苷酸重复

<220>
<221> misc_feature
<222> (374)..(436)
<223> 外显子 1B

<400> 111
acguaaccua cggugucccg cuaggaaaga gaggugcguc aaacagcgac aaguuccgcc 60

cacguaaaag augacgcuug gugugucagc cgucccugcu gcccgguugc uucucuuuug 120

ggggcggggu cuagcaagag cagguguggg uuuaggaggu guguguuuuu guuuuuccca 180

cccucucucc ccacuacuug cucucacagu acucgcugag ggugaacaag aaaagaccug 240

auaaagauua accagaagaa aacaaggagg gaaacaaccg cagccuguag caagcucugg 300

aacucaggag ucgcgcgcua ggggccgggg ccggggccgg ggcguggucg gggcgggccc 360

ggggggcggc ccggggcggg gcugcgguug cggugccugc gcccgcggcg gcggaggcgc 420

aggcgguggc gaguggguga gugaggaggc ggcauccugg cggguggcug uuugggguuc 480

ggcugccggg aagaggcgcg gguagaagcg ggggcucucc ucagagcucg acgcauuuuu 540

acuuucccuc ucauuucucu gaccgaagcu gggugucggg cuuucgccuc uagcgacugg 600

uggaauugcc ugcauccggg ccccgggcuu cccggcggcg gcggcggcgg cggcggcgca 660

gggacaaggg auggggaucu ggccucuucc uugcuuuccc gcccucagua cccgagcugu 720

cuccuucccg gggacccgcu gggagcgcug ccgcugcggg cucgagaaaa gggagccucg 780

gguacugaga ggccucgccu gggggaaggc cggagggugg gcggcgcgcg gcuucugcgg 840

accaagucgg gguucgcuag gaacccgaga cggucccugc cggcgaggag aucaugcggg 900

augagauggg gguguggaga cgccugcaca auuucagccc aagcuucuag agagugguga 960

ugacuugcau augagggcag caaugcaagu cggugugcuc cccauucugu gggacaugac 1020

cugguugcuu cacagcuccg agaugacaca gacuugcuua aaggaaguga cuauugugac 1080

uugggcauca cuugacugau gguaaucagu ugucuaaaga agugcacaga uuacaugucc 1140

gugugcucau uggguauc uggccgcguu gaacaccacc aggcuuugua uucagaaaca 1200

ggagggaggu ccugcacuuu cccaggaggg guggcccuuu cagaugcaau cgagauuguu 1260

[1113] aggcucuggg agaguaguug ccugguugug gcaguuggua aauuucuauu caaacaguug 1320

ccaugcacca guuguucaca acaaggguac guaaucuguc uggcauuacu ucuacuuuug 1380

uacaaaggau caaaaaaaaa aaagauacug uuaagauaug auuuuucuca gacuuuggga 1440

aacuuuuaac auaaucugug aauaucacag aaacaagacu aucauauagg ggauauuaau 1500

aaccuggagu cagaauacuu gaaauacggu gucauuugac acgggcauug uugucaccac 1560

cucugccaag gccugccacu uuaggaaaac ccugaaucag uuggaaacug cuacaugcug 1620

auaguacauc ugaaacaaga acgagaguaa uuaccacauu ccagauuguu cacuaagcca 1680

gcauuuaccu gcuccaggaa aaaauuacaa gcaccuuaug aaguugauaa aauauuuugu 1740

uuggcuaugu uggcacucca caauuugcuu ucagagaaac aaaguaaacc aaggaggacu 1800

ucuguuuuuc aagucugccc ucggguucua uucuacguua auuagauagu ucccaggagg 1860

acuagguuag ccuaccuauu gucugagaaa cuuggaacug ugagaaaugg ccagauagug 1920

auaugaacuu caccuuccag ucuucccuga uguugaagau ugagaaagug uugugaacuu 1980

ucugguacug uaaacaguuc acuguccuug aagugguccu gggcagcucc uguuguggaa 2040

aguggacggu uuaggauccu gcuucucuuu gggcugggag aaaauaaaca gcaugguuac 2100

aaguauugag agccagguug gagaaggugg cuuacaccug uaaugccaga gcuuugggag 2160

gcggaggcaa gaggaucacu ugaagccagg aguucaagcu caaccugggc aacguagacc 2220

cugucucuac aaaaaauuaa aaacuuagcc gggcguggug augugcaccu guaguccuag 2280

cuacuuggga ggcugaggca ggagggucau uugagcccaa gaguuugaag uuaccgagag 2340

cuaugauccu gccagugcau uccagccugg augacaaaac gagacccugu cucuaaaaaa 2400

caagaaguga gggcuuuaug auuguagaau uuucacuaca auagcagugg accaaccacc 2460

uuucuaaaua ccaaucaggg aagagauggu ugauuuuuua acagacguuu aaagaaaaag 2520

caaaaccuca aacuuagcac ucuacuaaca guuuuagcag auguuaauua auguaaucau 2580

gucugcaugu augggauuau uuccagaaag uguauuggga aaccucucau gaacccugug 2640

agcaagccac cgucucacuc aauuugaauc uuggcuuccc ucaaaagacu ggcuaauguu 2700

ugguaacucu cuggaguaga cagcacuaca uguacguaag auagguacau aaacaacuau 2760

ugguuuugag cugauuuuuu ucagcugcau uugcauguau ggauuuuucu caccaaagac 2820

gaugacuuca aguauuagua aaauaauugu acagcucucc ugauuauacu ucucugugac 2880

auuucauuuc ccaggcuauu ucuuuuggua ggauuuaaaa cuaagcaauu caguaugauc 2940

uuuguccuuc auuuucuuuc uuauucuuuu uguuuguuug uuuguuuguu uuuuucuuga 3000

ggcagagucu cucucugucg cccaggcugg agugcagugg cgccaucuca gcucauugca 3060

accucugcca ccuccggguu caagagauuc uccugccuca gccucccgag uagcugggau 3120

uacagguguc caccaccaca cccggcuaau uuuuuguauu uuuaguagag gugggguuuc 3180

accauguugg ccaggcuggu cuugagcucc ugaccucagg ugauccaccu gccucggccu 3240

[1114] accaaagagc ugggauaaca ggugugaccc accaugcccg gcccauuuuu uuuuucuuau 3300

ucuguuagga gugagagugu aacuagcagu auaauaguuc aauuuucaca acgugguaaa 3360

aguuucccua uaauucaauc agauuuugcu ccaggguuca guucuguuuu aggaaauacu 3420

uuuauuuuca guuuaaugau gaaauauuag aguuguaaua uugccuuuau gauuauccac 3480

cuuuuuaacc uaaaagaaug aaagaaaaau auguuugcaa uauaauuuua ugguuguaug 3540

uuaacuuaau ucauuauguu ggccuccagu uugcuguugu uaguuaugac agcaguagug 3600

ucauuaccau uucaauucag auuacauucc uauauuugau cauuguaaac ugacugcuua 3660

cauuguauua aaaacagugg auauuuuaaa gaagcuguac ggcuuauauc uagugcuguc 3720

ucuuaagacu auuaaauuga uacaacauau uuaaaaguaa auauuaccua aaugaauuuu 3780

ugaaauuaca aauacacgug uuaaaacugu cguuguguuc aaccauuucu guacauacuu 3840

agaguuaacu guuuugccag gcucuguaug ccuacucaua auaugauaaa agcacucauc 3900

uaaugcucug uaaauagaag ucagugcuuu ccaucagacu gaacucucuu gacaagaugu 3960

ggaugaaauu cuuuaaguaa aauuguuuac uuugucauac auuuacagau caaauguuag 4020

cucccaaagc aaucauaugg caaagauagg uauaucauag uuugccuauu agcugcuuug 4080

uauugcuauu auuauaaaua gacuucacag uuuuagacuu gcuuagguga aauugcaauu 4140

cuuuuuacuu ucagucuuag auaacaaguc uucaauuaua guacaaucac acauugcuua 4200

ggaaugcauc auuaggcgau uuugucauua ugcaaacauc auagagugua cuuacacaaa 4260

ccuagauagu auagccuuua uguaccuagg ccguauggua uagucuguug cuccuaggcc 4320

acaaaccugu acaacuguua cuguacugaa uacuauagac aguuguaaca cagugguaaa 4380

uauuuaucua aauauaugca aacagagaaa agguacagua aaaguauggu auaaaagaua 4440

augguauacc uguguaggcc acuuaccacg aauggagcuu gcaggacuag aaguugcucu 4500

gggugaguca gugagugagu ggugaauuaa ugugaaggcc uagaacacug uacaccacug 4560

uagacuauaa acacaguacg cugaagcuac accaaauuua ucuuaacagu uuuucuucaa 4620

uaaaaaauua uaacuuuuua acuuuguaaa cuuuuuaauu uuuuaacuuu uaaaauacuu 4680

agcuugaaac acaaauacau uguauagcua uacaaaaaua uuuuuucuuu guauccuuau 4740

ucuagaagcu uuuuucuauu uucuauuuua aauuuuuuuu uuuacuuguu agucguuuuu 4800

guuaaaaacu aaaacacaca cacuuucacc uaggcauaga caggauuagg aucaucagua 4860

ucacucccuu ccaccucacu gccuuccacc uccacaucuu gucccacugg aagguuuuua 4920

ggggcaauaa cacacaugua gcugucaccu augauaacag ugcuuucugu ugaauaccuc 4980

cugaaggacu ugccugaggc uguuuuacau uuaacuuaaa aaaaaaaaaa guagaaggag 5040

ugcacucuaa aauaacaaua aaaggcauag uauagugaau acauaaacca gcaauguagu 5100

aguuuauuau caaguguugu acacuguaau aauuguaugu gcuauacuuu aaauaacuug 5160

caaaauagua cuaagaccuu augaugguua cagugucacu aaggcaauag cauauuuuca 5220

[1115] gguccauugu aaucuaaugg gacuaccauc auauaugcag ucuaccauug acugaaacgu 5280

uacauggcac auaacuguau uugcaagaau gauuuguuuu acauuaauau cacauaggau 5340

guaccuuuuu agagugguau guuuaugugg auuaagaugu acaaguugag caaggggacc 5400

aagagcccug gguucugucu uggaugugag cguuuauguu cuucuccuca ugucuguuuu 5460

cucauuaaau ucaaaggcuu gaacgggccc uauuuagccc uucuguuuuc uacguguucu 5520

aaauaacuaa agcuuuuaaa uucuagccau uuaguguaga acucucuuug cagugaugaa 5580

augcuguauu gguuucuugg cuagcauauu aaauauuuuu aucuuugucu ugauacuuca 5640

augucguuuu aaacaucagg aucgggcuuc aguauucuca uaaccagaga guucacugag 5700

gauacaggac uguuugccca uuuuuuguua uggcuccaga cuugugguau uuccaugucu 5760

uuuuuuuuuu uuuuuuuuuu gaccuuuuag cggcuuuaaa guauuucugu uguuaggugu 5820

uguauuacuu uucuaagauu acuuaacaaa gcaccacaaa cugaguggcu uuaaacaaca 5880

gcaauuuauu cucucacaau ucuagaagcu agaaguccga aaucaaagug uugacagggg 5940

caugaucuuc aagagagaag acucuuuccu ugccucuucc uggcuucugg ugguuaccag 6000

caauccugag uguuccuuuc uugccuugua guuucaacaa uccaguaucu gccuuuuguc 6060

uucacauggc ugucuaccau uugucucugu gucuccaaau cucucuccuu auaaacacag 6120

caguuauugg auuaggcccc acucuaaucc aguaugaccc cauuuuaaca ugauuacacu 6180

uauuucuaga uaaggucaca uucacguaca ccaaggguua ggaauugaac auaucuuuuu 6240

```
gggggacaca auucaaccca caagugucag ucucuagcug agccuuuccc uuccuguuuu    6300

ucuccuuuuu aguugcuaug gguuaggggc caaaucucca gucauacuag aauugcacau    6360

ggacuggaua uuugggaaua cugcgggucu auucuaugag cuuuaguaug uaacauuuaa    6420

uaucagugua aagaagcccu uuuuuaaguu auuucuuuga auuucuaaau guaugcccug    6480

aauauaagua acaaguuacc augucuugua aaaugaucau aucaacaaac auuuaaugug    6540

caccuacugu gcuaguugaa ugucuuuauc cugauaggag auaacaggau uccacaucuu    6600

ugacuuaaga ggacaaacca aauaugucua aaucauuugg gguuuugaug gauaucuuua    6660

aauugcugaa ccuaaucauu gguuucauau gucauuguuu ag                      6702
```

<210> 112
<211> 6702
<212> RNA
<213> 人工序列

<220>
<223> 合成

<220>
<221> misc_feature
<222> (6267)..(6329)
<223> 外显子 1B

[1116]

<220>
<221> misc_feature
<222> (6365)..(6382)
<223> 3x 六核苷酸重复

<220>
<221> misc_feature
<222> (6545)..(6702)
<223> 外显子 1A

<400> 112

```
cuaaacaaug acauaugaaa ccaaugauua gguucagcaa uuuaaagaua uccaucaaaa      60

ccccaaauga uuuagacaua uuugguuugu ccucuuaagu caaagaugug gaauccuguu     120

aucuccuauc aggauaaaga cauucaacua gcacaguagg ugcacauuaa auguuuguug     180

auaugaucau uuuacaagac augguaacuu guuacuuaua uucagggcau acauuuagaa     240

auucaaagaa auaacuuaaa aaagggcuuc uuuacacuga uauuaaaugu uacauacuaa     300

agcucauaga auagacccgc aguauuccca aauauccagu ccaugugcaa uucuaguaug     360

acuggagauu uggccccuaa cccauagcaa cuaaaaagga gaaaaacagg aagggaaagg     420

cucagcuaga gacugacacu ugugggguga auuguguccc ccaaaaagau auguucaauu     480

ccuaacccuu gguguacgug aaugugaccu uaucuagaaa uaaguguaau cauguuaaaa     540

uggggucaua cuggauuaga gugggggccua auccaauaac ugcuguguuu auaaggagag     600

agauuuggag acacagagac aaaugguaga cagccaugug aagacaaaag gcagauacug     660

gauuguugaa acuacaaggc aagaaaggaa cacucaggau ugcugguaac caccagaagc     720
```

caggaagagg caaggaaaga gucuucucuc uugaagauca ugccccuguc aacacuuuga 780

uuucggacuu cuagcuucua gaauugugag agaauaaauu gcuguuguuu aaagccacuc 840

aguuuguggu gcuuuguuaa guaaucuuag aaaaguaaua caacaccuaa caacagaaau 900

acuuuaaagc cgcuaaaagg ucaaaaaaaa aaaaaaaaaa aaagacaugg aaauaccaca 960

agucuggagc cauaacaaaa aaugggcaaa caguccugua uccucaguga acucucuggu 1020

uaugagaaua cugaagcccg auccugaugu uuaaaacgac auugaaguau caagacaaag 1080

auaaaaauau uuaauaugcu agccaagaaa ccaauacagc auuucaucac ugcaaagaga 1140

guucuacacu aaauggcuag aauuuaaaag cuuuaguuau uuagaacacg uagaaaacag 1200

aagggcuaaa uagggcccgu ucaagccuuu gaauuuaaug agaaaacaga caugaggaga 1260

agaacauaaa cgcucacauc caagacagaa cccagggcuc uugguccccu ugcucaacuu 1320

guacaucuua auccacauaa acauaccacu cuaaaaaggu acauccuaug ugauauuaau 1380

guaaaacaaa ucauucuugc aaauacaguu augugccaug uaacguuuca gucaauggua 1440

gacugcauau augaugguag ucccauuaga uuacaaugga ccugaaaaua ugcuauugcc 1500

uuagugacac uguaaccauc auaaggucuu aguacuauuu ugcaaguuau uuaaaguaua 1560

gcacauacaa uuauuacagu guacaacacu ugauaauaaa cuacuacauu gcugguuuau 1620

guauucacua uacuaugccu uuuauuguua uuuuagagug cacuccuucu acuuuuuuuu 1680

[1117] uuuuuaaguu aaauguaaaa cagccucagg caaguccuuc aggagguauu caacagaaag 1740

cacuguuauc auaggugaca gcuacaugug uguuauugcc ccuaaaaacc uuccaguggg 1800

acaagaugug gagguggaag gcagugaggu ggaagggagu gauacugaug auccuaaucc 1860

ugucuaugcc uaggugaaag ugugugguu uuaguuuuua acaaaaacga cuaacaagua 1920

aaaaaaaaaa uuuaaaauag aaaauagaaa aaagcuucua gaauaaggau acaaagaaaa 1980

aauauuuuug uauagcuaua caauguauuu guguuucaag cuaaguauuu uaaaaguuaa 2040

aaaauuaaaa aguuuacaaa guuaaaaagu uauaauuuuu uauugaagaa aaacuguuaa 2100

gauaaauuug guguagcuuc agcguacugu guuuauaguc uacaguggug uacaguguuc 2160

uaggccuuca cauuaauuca ccacucacuc acugacucac ccagagcaac uucuaguccu 2220

gcaagcucca uucgugguaa guggccuaca cagguauacc auuaucuuuu auaccauacu 2280

uuuacuguac cuuuucucug uuugcauaua uuuagauaaa uauuuaccac uguguuacaa 2340

cugucuauag uauucaguac aguaacaguu guacagguuu guggccuagg agcaacagac 2400

uauaccauac ggccuaggua cauaaaggcu auacuaucua gguuugugua aguacacucu 2460

augauguuug cauaaugaca aaaucgccua augaugcauu ccuaagcaau gugugauugu 2520

acuauaauug aagacuuguu aucuaagacu gaaaguaaaa agaauugcaa uuucaccuaa 2580

gcaagucuaa aacugugaag ucuauuuaua auaauagcaa uacaaagcag cuaauaggca 2640

aacuaugaua uaccuaucuu ugccauauga uugcuuuggg agcuaacauu ugaucuguaa 2700

auguaugaca aaguaaacaa uuuuacuuaa agaauuucau ccacaucuug ucaagagagu 2760

ucagucugau ggaaagcacu gacuucuauu uacagagcau uagaugagug cuuuuaucau 2820

auuaugagua ggcauacaga gccuggcaaa acaguuaacu cuaaguaugu acagaaaugg 2880

uugaacacaa cgacaguuuu aacacgugua uuuguaauuu caaaaauuca uuuagguaau 2940

auuuacuuuu aaauauguug uaucaauuua auagucuuaa gagacagcac uagauauaag 3000

ccguacagcu ucuuuaaaau auccacuguu uuuaauacaa uguaagcagu caguuuacaa 3060

ugaucaaaua uaggaaugua aucugaauug aaaugguaau gacacuacug cugucauaac 3120

uaacaacagc aaacuggagg ccaacauaau gaauuaaguu aacauacaac cauaaaauua 3180

uauugcaaac auauuuuucu uucauucuuu uagguuaaaa agguggauaa ucauaaaggc 3240

aauauuacaa cucuaauauu ucaucauuaa acugaaaaua aaaguauuuc cuaaaacaga 3300

acugaacccu ggagcaaaau cugauugaau uauagggaaa cuuuuaccac guugugaaaa 3360

uugaacuauu auacugcuag uuacacucuc acuccuaaca gaauaagaaa aaaaaaaugg 3420

gccgggcaug gugggucaca ccuguuaucc cagcucuuug guaggccgag gcagguggau 3480

caccugaggu caggagcuca agaccagccu ggccaacaug gugaaacccc accucuacua 3540

aaaauacaaa aaauuagccg ggugguggugg uggacaccug uaaucccagc uacucgggag 3600

gcugaggcag gagaaucucu ugaacccgga gguggcagag guugcaauga gcugagaugg 3660

[1118] cgccacugca cuccagccug ggcgacagag agagacucug ccucaagaaa aaaacaaaca 3720

aacaaacaaa caaaaagaau aagaaagaaa augaaggaca aagaucauac ugaauugcuu 3780

aguuuuaaau ccuaccaaaa gaaauagccu gggaaaugaa augucacaga gaaguauaau 3840

caggagagcu guacaauuau uuuacuaaua cuugaaguca ucgucuuugg ugagaaaaau 3900

ccauacaugc aaaugcagcu gaaaaaaauc agcucaaaac caauaguugu uuauguaccu 3960

aucuuacgua cauguagugc ugucuacucc agagaguuac caaacauuag ccagucuuuu 4020

gagggaagcc aagauucaaa uugagugaga cgguggcuug cucacagggu ucaugagagg 4080

uuucccaaua cacuuucugg aaauaauccc auacaugcag acaugauuac auuaauuaac 4140

aucugcuaaa acuguuagua gagugcuaag uuugagguuu ugcuuuuucu uuaaacgucu 4200

guuaaaaaau caaccaucuc uucccugauu gguauuuaga aaggugguug guccacugcu 4260

auuguaguga aaauucuaca aucauaaagc ccucacuucu uguuuuuuag agacaggguc 4320

ucguuuuguc auccaggcug gaaugcacug gcaggaucau agcucucggu aacuucaaac 4380

ucuugggcuc aaaugacccu ccugccucag ccucccaagu agcuaggacu acaggugcac 4440

aucaccacgc ccggcuaagu uuuuaauuuu uuguagagac agggucuacg uugcccaggu 4500

ugagcuugaa cuccuggcuu caagugaucc ucuugccucc gccucccaaa gcucuggcau 4560

uacaggugua agccaccuuc uccaaccugg cucucaauac uuguaaccau gcuguuuauu 4620

uucucccagc ccaaagagaa gcaggauccu aaaccgucca cuuuccacaa caggagcugc 4680

ccaggaccac uucaaggaca gugaacuguu uacaguacca gaaaguucac aacacuuucu 4740

caaucuucaa caucagggaa gacuggaagg ugaaguucau aucacuaucu ggccauuucu 4800

cacaguucca aguuucucag acaauaggua ggcuaaccua guccuccugg gaacuaucua 4860

auuaacguag aauagaaccc gagggcagac uugaaaaaca gaaguccucc uugguuuacu 4920

uuguuucucu gaaagcaaau uguggagugc caacauagcc aaacaaaaua uuuuaucaac 4980

uucauaaggu gcuuguaauu uuuuccugga gcagguaaau gcuggcuuag ugaacaaucu 5040

ggaauguggu aauuacucuc guucuuguuu cagauguacu aucagcaugu agcaguuucc 5100

aacugauuca ggguuuuccu aaaguggcag gccuuggcag agguggugac aacaaugccc 5160

gugucaaaug acaccguauu ucaaguauuc ugacuccagg uuauuaauau ccccuauaug 5220

auagucuugu uucugugaua uucacagauu auguuaaaag uuucccaaag ucugagaaaa 5280

aucauaucuu aacaguaucu uuuuuuuuuu ugauccuuug uacaaaagua gaaguaaugc 5340

cagacagauu acguacccuu guugugaaca acuggugcau ggcaacuguu ugaauagaaa 5400

uuuaccaacu gccacaacca ggcaacuacu cucccagagc cuaacaaucu cgauugcauc 5460

ugaaagggcc accccuccug ggaaagugca ggaccucccu ccuguuucug aauacaaagc 5520

cugguggugu ucaacgcggc cagauagacc caaugagcac acggacaugu aaucugugca 5580

cuucuuuaga caacugauua ccaucaguca agugaugccc aagucacaau agucacuucc 5640

[1119] uuuaagcaag ucugugucau cucggagcug ugaagcaacc aggucauguc ccacagaaug 5700

gggagcacac cgacuugcau ugcugcccuc auaugcaagu caucaccacu cucuagaagc 5760

uuggggcugaa auugugcagg cgucuccaca cccccaucuc aucccgcaug aucuccucgc 5820

cggcagggac cgucucgggu uccuagcgaa ccccgacuug guccgcagaa gccgcgcgcc 5880

gcccacccuc cggccuuccc ccaggcgagg ccucucagua cccgaggcuc ccuuuucucg 5940

agcccgcagc ggcagcgcuc ccagcggguc cccgggaagg agacagcucg gguacugagg 6000

gcgggaaagc aaggaagagg ccagaucccc aucccuuguc ccugcgccgc cgccgccgcc 6060

gccgccgccg ggaagcccgg ggcccggaug caggcaauuc caccagucgc uagaggcgaa 6120

agcccgacac ccagcuucgg ucagagaaau gagagggaaa guaaaaaugc gucgagcucu 6180

gaggagagcc cccgcuucua cccgcgccuc uucccggcag ccgaacccca aacagccacc 6240

cgccaggaug ccgccuccuc acucacccac ucgccaccgc cugcgccucc gccgccgcgg 6300

gcgcaggcac cgcaaccgca gccccgcccc gggcccgccc ccgggcccgc cccgaccacg 6360

ccccggcccc ggccccggcc ccuagcgcgc gacuccugag uuccagagcu ugcuacaggc 6420

ugcgguuguu ucccuccuug uuuucuucug guuaaucuuu aucaggucuu uucuuguuca 6480

cccucagcga guacugugag agcaaguagu ggggagagag gguggggaaaa acaaaaacac 6540

acaccuccua aacccacacc ugcucuugcu agaccccgcc cccaaaagag aagcaaccgg 6600

gcagcaggga cggcugacac accaagcguc aucuuuuacg ugggcggaac uugucgcugu 6660

uugacgcacc ucucuuuccu agcgggacac cguagguuac gu 6702

<210> 113
<211> 215
<212> RNA
<213> 人工序列

<220>
<223> 合成

<220>
<221> misc_feature
<222> (163)..(180)
<223> 3x 六核苷酸重复

<400> 113
gugugugguu uuguuuuucc cacccucucu ccccacuacu ugcucucaca guacucgcug 60

agggugaaca agaaaagacc ugauaaagau uaaccagaag aaaacaagga gggaaacaac 120

cgcagccugu agcaagcucu ggaacucagg agucgcgcgc uaggggccgg ggccggggcc 180

ggggcguggu cggggcgggc ccgggggcgg gcccg 215

<210> 114
<211> 215
<212> RNA
<213> 人工序列

[1120]

<220>
<223> 合成

<220>
<221> misc_feature
<222> (36)..(53)
<223> 3x 六核苷酸重复

<400> 114
cgggcccgcc cccgggcccg ccccgaccac gccccggccc cggccccggc cccuagcgcg 60

cgacuccuga guuccagagc uugcuacagg cugcgguugu uucccuccuu guuuucuucu 120

gguuaaucuu uaucaggucu uuucuuguuc acccucagcg aguacuguga gagcaaguag 180

uggggagaga ggguggggaaa aacaaaaaca cacac 215

<210> 115
<211> 162
<212> RNA
<213> 人工序列

<220>
<223> 合成

<400> 115
gugugugguu uuguuuuucc cacccucucu ccccacuacu ugcucucaca guacucgcug 60

agggugaaca agaaaagacc ugauaaagau uaaccagaag aaaacaagga gggaaacaac 120

cgcagccugu agcaagcucu ggaacucagg agucgcgcgc ua 162

<210> 116
<211> 162
<212> RNA
<213> 人工序列

<220>
<223> 合成

<400> 116
uagcgcgcga cuccugaguu ccagagcuug cuacaggcug cgguuguuuc ccuccuuguu 60

uucuucuggu uaaucuuuau caggucuuuu cuuguucacc cucagcgagu acugugagag 120

caaguagugg ggagagaggg ugggaaaaac aaaaacacac ac 162

<210> 117
<211> 36
<212> RNA
<213> 人工序列

<220>
<223> 合成

<400> 117
caagaaaaga ccugauaaag auuaaccaga agaaaa 36

[1121]

<210> 118
<211> 36
<212> RNA
<213> 人工序列

<220>
<223> 合成

<400> 118
uuuucuucug guuaaucuuu aucaggucuu uucuug 36

SEQ ID NO:121
>NM_001256054.2 智人 (Homo sapiens) C9orf72-SMCR8 复合物亚基 (C9orf72), 转录本变体 3, mRNA
ACGTAACCTACGGTGTCCCGCTAGGAAAGAGAGGTGCGTCAAACAGCGACAAGTTCCGCCCACGTAAAAG
ATGACGCTTGGTGTGTCAGCCGTCCCTGCTGCCCGGTTGCTTCTCTTTTGGGGGCGGGGTCTAGCAAGAG
CAGGTGTGGGTTTAGGAGATATCTCCGGAGCATTTGGATAATGTGACAGTTGGAATGCAGTGATGTCGAC
TCTTTGCCCACCGCCATCTCCAGCTGTTGCCAAGACAGAGATTGCTTTAAGTGGCAAATCACCTTTATTA
GCAGCTACTTTTGCTTACTGGGACAATATTCTTGGTCCTAGAGTAAGGCACATTTGGGCTCCAAAGACAG
AACAGGTACTTCTCAGTGATGGAGAAATAACTTTTCTTGCCAACCACACTCTAAATGGAGAAATCCTTCG
AAATGCAGAGAGTGGTGCTATAGATGTAAAGTTTTTTGTCTTGTCTGAAAAGGGAGTGATTATTGTTTCA
TTAATCTTTGATGGAAACTGGAATGGGGATCGCAGCACATATGGACTATCAATTATACTTCCACAGACAG
AACTTAGTTTCTACCTCCCACTTCATAGAGTGTGTGTTGATAGATTAACACATATAATCCGGAAAGGAAG
AATATGGATGCATAAGGAAAGACAAGAAAATGTCCAGAAGATTATCTTAGAAGGCACAGAGAGAATGGAA
GATCAGGGTCAGAGTATTATTCCAATGCTTACTGGAGAAGTGATTCCTGTAATGGAACTGCTTTCATCTA
TGAAATCACACAGTGTTCCTGAAGAAATAGATATAGCTGATACAGTACTCAATGATGATGATATTGGTGA
CAGCTGTCATGAAGGCTTTCTTCTCAATGCCATCAGCTCACACTTGCAAACCTGTGGCTGTTCCGTTGTA
GTAGGTAGCAGTGCAGAGAAAGTAAATAAGATAGTCAGAACATTATGCCTTTTTCTGACTCCAGCAGAGA
GAAAATGCTCCAGGTTATGTGAAGCAGAATCATCATTTAAATATGAGTCAGGGCTCTTTGTACAAGGCCT
GCTAAAGGATTCAACTGGAAGCTTTGTGCTGCCTTTCCGGCAAGTCATGTATGCTCCATATCCCACCACA
CACATAGATGTGGATGTCAATACTGTGAAGCAGATGCCACCCTGTCATGAACATATTTATAATCAGCGTA
GATACATGAGATCCGAGCTGACAGCCTTCTGGAGAGCCACTTCAGAAGAAGACATGGCTCAGGATACGAT
CATCTACACTGACGAAAGCTTTACTCCTGATTTGAATATTTTTCAAGATGTCTTACACAGAGACACTCTA
GTGAAAGCCTTCCTGGATCAGGTCTTTCAGCTGAAACCTGGCTTATCTCTCAGAAGTACTTTCCTTGCAC

AGTTTCTACTTGTCCTTCACAGAAAAGCCTTGACACTAATAAAATATATAGAAGACGATACGCAGAAGGG
AAAAAAGCCCTTTAAATCTCTTCGGAACCTGAAGATAGACCTTGATTTAACAGCAGAGGGCGATCTTAAC
ATAATAATGGCTCTGGCTGAGAAAATTAAACCAGGCCTACACTCTTTTATCTTTGGAAGACCTTTCTACA
CTAGTGTGCAAGAACGAGATGTTCTAATGACTTTTTAAATGTGTAACTTAATAAGCCTATTCCATCACAA
TCATGATCGCTGGTAAAGTAGCTCAGTGGTGTGGGGAAACGTTCCCCTGGATCATACTCCAGAATTCTGC
TCTCAGCAATTGCAGTTAAGTAAGTTACACTACAGTTCTCACAAGAGCCTGTGAGGGGATGTCAGGTGCA
TCATTACATTGGGTGTCTCTTTTCCTAGATTTATGCTTTTGGGATACAGACCTATGTTTACAATATAATA
AATATTATTGCTATCTTTTAAAGATATAATAATAGGATGTAAACTTGACCACAACTACTGTTTTTTTGAA
ATACATGATTCATGGTTTACATGTGTCAAGGTGAAATCTGAGTTGGCTTTTACAGATAGTTGACTTTCTA
TCTTTTGGCATTCTTTGGTGTGTAGAATTACTGTAATACTTCTGCAATCAACTGAAAACTAGAGCCTTTA
AATGATTTCAATTCCACAGAAAGAAAGTGAGCTTGAACATAGGATGAGCTTTAGAAAGAAAATTGATCAA
GCAGATGTTTAATTGGAATTGATTATTAGATCCTACTTTGTGGATTTAGTCCCTGGGATTCAGTCTGTAG
AAATGTCTAATAGTTCTCTATAGTCCTTGTTCCTGGTGAACCACAGTTAGGGTGTTTTGTTTATTTTATT
GTTCTTGCTATTGTTGATATTCTATGTAGTTGAGCTCTGTAAAAGGAAATTGTATTTTATGTTTTAGTAA
TTGTTGCCAACTTTTTAAATTAATTTTCATTATTTTTGAGCCAAATTGAAATGTGCACCTCCTGTGCCTT
TTTTCTCCTTAGAAAATCTAATTACTTGGAACAAGTTCAGATTTCACTGGTCAGTCATTTTCATCTTGTT
TTCTTCTTGCTAAGTCTTACCATGTACCTGCTTTGGCAATCATTGCAACTCTGAGATTATAAAATGCCTT
AGAGAATATACTAACTAATAAGATCTTTTTTTCAGAAACAGAAAATAGTTCCTTGAGTACTTCCTTCTTG
CATTTCTGCCTATGTTTTTGAAGTTGTTGCTGTTTGCCTGCAATAGGCTATAAGGAATAGCAGGAGAAAT
TTTACTGAAGTGCTGTTTTCCTAGGTGCTACTTTGGCAGAGCTAAGTTATCTTTTGTTTTCTTAATGCGT
TTGGACCATTTTGCTGGCTATAAAATAACTGATTAATATAATTCTAACACAATGTTGACATTGTAGTTAC
ACAAACACAAATAAATATTTTATTTAAAATTCTGGAAGTAATATAAAAGGGAAAATATATTTATAAGAAA
GGGATAAAGGTAATAGAGCCCTTCTGCCCCCCACCCACCAAATTTACACAACAAAATGACATGTTCGAAT
GTGAAAGGTCATAATAGCTTTCCCATCATGAATCAGAAAGATGTGGACAGCTTGATGTTTTAGACAACCA
CTGAACTAGATGACTGTTGTACTGTAGCTCAGTCATTTAAAAAATATATAAATACTACCTTGTAGTGTCC
CATACTGTGTTTTTTACATGGTAGATTCTTATTTAAGTGCTAACTGGTTATTTTCTTTGGCTGGTTTATT
GTACTGTTATACAGAATGTAAGTTGTACAGTGAAATAAGTTATTAAAGCATGTGTAAACATTGTTATATA
TCTTTTCTCCTAAATGGAGAATTTTGAATAAAATATATTTGAAATTTTAAAAAAAAAAAAAAAAAA

SEQ ID NO:122

>XM_005581570.2 PREDICTED: 食蟹猴(Macaca fascicularis) 染色体 15 开放阅读框, huma C9orf72 (C15H9orf72), 转录本变体 X2, mRNA

[1122]

ACGTAACCTACGGTGTCCCGCTAGGAAAGAGAGGCGCGTCAAACAGCGACAAGTTCCGCCCACGTAAAAG
ATGACGCTTGGTGCGTCAGCCGTCCCTGCTGCCCGGTTCCTTCTCTCTGGGGGCGGGGCCTGGCTAGAGC
AGGTGTGGGTTTAGGAGATATCTCAGGAGCATTTGGATAATGTGACAGTTGGAATGCAGTGATGTCGACT
CTTTGCCCACCGCCATCTCCAGCTGTTGCCAAGACAGAGATTGCTTTAAGTGGTGAATCACCTTTATTAG
CAGCTACTTTTGCTTACTGGGACAATATTCTTGGTCCTAGAGTAAGGCACATTTGGGCTCCAAAGACAGA
ACAGGTACTTCTCAGTGACGGAGAAATAACTTTTCTTGCCAACCACACTCTAAATGGAGAAATCCTTCGA
AATGCAGAGAGTGGTGCTATAGATGTAAAGTTTTTTGTCTTGTCTGAAAAGGGAGTGATTATTGTTTCAT
TAATCTTTGATGGAAACTGGAATGGGGATCGCAGCACATACGGACTATCAATTATACTTCCACAGACAGA
ACTTAGTTTCTACCTCCCACTTCATAGAGTGTGTGTTGATAGATTAACACATATAATCCGGAAAGGAAGA
ATATGGATGCATAAGGAAAGACAAGAAAATGTCCAGAAGATTATCTTAGAAGGCACAGAGAGAATGGAAG
ATCAGGGTCAGAGTATTATTCCAATGCTTACTGGAGAAGTGATTCCTGTAATGGAACTGCTTTCATCTAT
GAAATCACACAGTGTTCCTGAAGAAATAGATATAGCTGATACAGTACTCAATGATGATGATATTGGTGAC
AGTTGTCATGAAGGCTTTCTTCTCAATGCCATCAGCTCACACTTGCAAACCTGTGGCTGTTCCGTTGTAG
TAGGTAGCAGTGCAGAGAAAGTAAATAAGATAGTCAGAACATTATGCCTTTTTCTGACTGCAGCAGAGAG
AAAATGCTCCAGGTTATGTGAAGCAGAATCATCATTTAAATATGAGTCAGGGCTCTTTGTACAGGGCCTG
CTAAAGGATTCAACTGGAAGCTTTGTGCTGCCTTTCCGGCAAGTCATGTATGCTCCATATCCCACCACAC
ACATAGATGTGGATGTCAATACTGTGAAGCAGATGCCACCCTGTCATGAACATATTTATAATCAGCGTAG
ATACATGAGATCCGAGCTGACAGCCTTCTGGAGAGCCACTTCAGAAGAAGACATGGCTCAGGATACGATC
ATCTACACTGACGAAAGCTTTACTCCTGATTTGAATATTTTTCAAGATGTCTTACACAGAGACACTCTAG
TGAAAGCCTTCCTGGATCAGGTCTTTCAGCTGAAACCTGGCTTATCTCTCAGGAGTACTTTCCTTGCACA
GTTTTTACTTGTCCTTCACAGAAAAGCCTTGACACTAATAAAATATATAGAAGATGATACGCAGAAGGGA
AAAAAGCCCTTTAAATCTCTTCGGAACCTGAAGATAGACCTTGATTTAACAGCAGAGGGCGATCTTAACA
TAATAATGGCTCTGGCTGAGAAAATTAAACCAGGCCTACACTCTTTTATCTTTGGAAGACCTTTCTACAC
TAGTGTACAAGAACGAGATGTTCTAATGACTTTTTAAATGTGTAACTTAATAAGCCTATTCCATCACAAT
CGTGATCGCTGCTAAAGTAGCTCGGTGGTGTGGGGAAACATTCCCCTGGATCATACTCCAGAGCTCTGCT
CGGCAGTTGCAGTTAAGTTAGTTACACTACAGTTCTCACAAGAGTCTGTGAGGGGATGTCAGGTGCATCA
TTACATTGGATGTCTCTTTTCCTAGATTTATGCTTTTGGGATACAGACCTATGTTTACAATATAATAGGT
ATTATTGCTGTCTTTTAAATATATAATAATAGGATATAAACTTGACCACAACTGCTGTTTTTTTGAAATA
TATGATTCATGGTTTACATGTATTAAGGTGAAATCCGAGTTCGCTTTTACAGATATTAGTTGACTTTCTA
TCTTTTGGCATTCTTTGGTGTGTGGAATTACTGTAATACTTCTGCAATCAACTGAAAATTAGAGCCTTTA
AATGATTTCAGTTCCACAGAAAGAAAGTGAGCTTCAACATAGGATAAGCTTTAGAAAGAGAATTGATCAA
GCAGATGTTTAATTGGAATTGATTATTAGATCCTGCTTTGTGGATTTAGCCCTCGGGATTCAGTCTGTAG
AAATGTCTGATAGTTCTCTATAGTCCCTGCTCATGGTGAACCACAGTTAGGATGTTTTGTTTGTTTTATT
GTTGTTGCTATTGTTGATGTTCTATATAGTTGAGCTCTGTAAAAGGAAATTGTATTTTATGTTTTAGTAG

TTGTTGCCAACTTTTTAAATTAATTTTCATTATTTTTGAGCCAAATTGAAATGTGCACCTCCTGTGCCTT
TTTTTTCCTTGGAAAATCGAATTACTTGGAAGAAGTTCAGATTTCACTGGTCAGTCGTTTTCATCTTGTT
TTCTTCTTGCAGAGTCTTACCATGTACCTGCTTTGGCAATCATTGTAACTCTGAGATTATAAAATGCATT
AGAGAATATATTAACTAATAAGATCTTTTTTTTCAGGAACAGAAAATAGTTCCTTGAGTACTTCCTTCTT
ACATTTCTGCCCATGTTTTTGAAGTTGTTGCCATTTGCCTGCAATAGGCTATAAGGAATAGCAGGAGAAA
TTTTACTGAAGTGCTATTTTTCTAGGTGCTACTTTGGCAGAGCTAAGTGGTCTGTTTCTTTTGTTTCCTT
AATGCGTTTGGACCATTTTGCTGGCTGTAAAATAACTGATTAATATAATTCTAACACAATATTGACATTG
TAGTGTACACAAACACAAATATTTTATTTAAAACTGGAAGTAACATAAAAGGGAAAATATATTTATAAGA
AAGGAATAAAGGTAATAGAGCTCTTCTGTCCCCCAGCCACCAAATTTACACAACAAAATGATATGTTCTA
ATGTGAAAGGTCATAATAGCTTTCCCATCATTAATCAGAAAGATGTGGCAGCTTGATTTTTCAGACAACC
CCTGAACTAGATGACTGTTGTACTGTAGCTCAGTCATTTAAAAAATATATAAATACTATCTCGTAGTGTC
CCATACTATGTTTTTTACATGATAGATTCTTATTTAAGTGCTACCTGGTTATTTTCTTTGGCTGGTTTAT
TGTACTGTTATATAGAATGTAAGTTGTACAGTGAAATAAGTTATTAAAGCATGTGTAAACATTGTTATAT
ATCTTTTCTCCTAGATGGAGAATTTTGAATAAAATATATTTGAAATTTT

SEQ ID NO:123

>NM_001081343.2 小鼠(Mus musculus) C9orf72, C9orf72-SMCR8复合物 (C9orf72)成员, 转录本变体 1, mRNA

[1123]

GCGGTTGCGGTCCCTGCGCCGGCGGTGAAGGCGCAGCAGCGGCGAGTGGCTATTGCAAGCGTTCGGATAA
TGTGAGACCTGGAATGCAGTGAGACCTGGGATGCAGGGATGTCGACTATCTGCCCCCCACCATCTCCTGC
TGTTGCCAAGACAGAGATTGCTTTAAGTGGTGAATCACCCTTGTTGGCGGCTACCTTTGCTTACTGGGAT
AATATTCTTGGTCCTAGAGTAAGGCATATTTGGGCTCCAAAGACAGACCAAGTGCTTCTCAGTGATGGAG
AAATAACTTTTCTTGCCAACCACACTCTAAATGGAGAAATTCTTCGAAATGCAGAGAGTGGGGCTATAGA
TGTAAAATTTTTTGTCTTATCTGAAAAAGGGGTAATTATTGTTTCATTAATCTTCGACGGAAACTGGAAT
GGAGATCGGAGCACTTATGGACTATCAATTATACTGCCGCAGACAGAGCTGAGCTTCTACCTCCCACTTC
ACAGAGTGTGTGTTGACAGGCTAACACACATTATTCGAAAAGGAAGAATATGGATGCATAAGGAAAGACA
AGAAAATGTCCAGAAAATTGTCTTGGAAGGCACAGAGAGGATGGAAGATCAGGGTCAGAGTATCATTCCC
ATGCTTACTGGGGAAGTCATTCCTGTAATGGAGCTGCTTGCATCTATGAAATCCCACAGTGTTCCTGAAG
ACATTGATATAGCTGATACAGTGCTCAATGATGATGACATTGGTGACAGCTGTCACGAAGGCTTTCTTCT
CAATGCCATCAGCTCACACCTGCAGACCTGTGGCTGTTCCGTTGTAGTTGGCAGCAGTGCAGAGAAAGTA
AATAAGATAGTAAGAACGCTGTGCCTTTTTCTGACACCAGCAGAGAGGAAATGCTCCAGGCTGTGTGAAG
CAGAATCGTCCTTTAAGTACGAATCGGGACTCTTTGTGCAAGGCTTGCTAAAGGATGCAACAGGCAGTTT
TGTCCTACCCTTCCGGCAAGTTATGTATGCCCCGTACCCCACCACGCACATTGATGTGGATGTCAACACT
GTCAAGCAGATGCCACCGTGTCATGAACATATTTATAATCAACGCAGATACATGAGGTCAGAGCTGACAG
CCTTCTGGAGGGCAACTTCAGAAGAGGACATGGCGCAGGACACCATCATCTACACAGATGAGAGCTTCAC
TCCTGATTTGAATATTTTCCAAGATGTCTTACACAGAGACACTCTAGTGAAAGCCTTCCTGGATCAGGTC
TTCCATTTGAAGCCTGGCCTGTCTCTCAGGAGTACTTTCCTTGCACAGTTCCTCCTCATTCTTCACAGAA
AAGCCTTGACACTAATCAAGTACATCGAGGATGATACGCAGAAGGGGAAAAAGCCCTTTAAGTCTCTTCG
GAACCTGAAGATAGATCTTGATTTAACAGCAGAGGGCGATCTTAACATAATAATGGCTCTAGCTGAGAAA
ATTAAGCCAGGCCTACACTCTTTCATCTTTGGGAGACCTTTCTACACTAGTGTACAAGAACGTGATGTTC
TAATGACCTTTTGACCGTGTGGTTTGCTGTGTCTGTCTCTTCACAGTCACACCTGCTGTTACAGTGTCTC
AGCAGTGTGTGGGCACATCCTTCCTCCCGAGTCCTGCTGCAGGACAGGGTACACTACACTTGTCAGTAGA
AGTCTGTACCTGATGTCAGGTGCATCGTTACAGTGAATGACTCTTCCTAGAATAGATGTACTCTTTTAGG
GCCTTATGTTTACAATTATCCTAAGTACTATTGCTGTCTTTTAAAGATATGAATGATGGAATATACACTT
GACCATAACTGCTGATTGGTTTTTTGTTTTGTTTTGTTTGTTTTCTTGGAAACTTATGATTCCTGGTTTA
CATGTACCACACTGAAACCCTCGTTAGCTTTACAGATAAAGTGTGAGTTGACTTCCTGCCCCTCTGTGTT
CTGTGGTATGTCCGATTACTTCTGCCACAGCTAAACATTAGAGCATTTAAAGTTTGCAGTTCCTCAGAAA
GGAACTTAGTCTGACTACAGATTAGTTCTTGAGAGAAGACACTGATAGGGCAGAGCTGTAGGTGAAATCA
GTTGTTAGCCCTTCCTTTATAGACGTAGTCCTTCAGATTCGGTCTGTACAGAAATGCCGAGGGGTCATGC
ATGGGCCCTGAGTATCGTGACCTGTGACAAGTTTTTTGTTGGTTTATTGTAGTTCTGTCAAAGAAAGTGG
CATTTGTTTTTATAATTGTTGCCAACTTTTAAGGTTAATTTTCATTATTTTTGAGCCGAATTAAAATGCG
CACCTCCTGTGCCTTTCCCAATCTTGGAAAATATAATTTCTTGGCAGAGGGTCAGATTTCAGGGCCCAGT
CACTTTCATCTGACCACCCTTTGCACGGCTGCCGTGTGCCTGGCTTAGATTAGAAGTCCTTGTTAAGTAT
GTCAGAGTACATTCGCTGATAAGATCTTTGAAGAGCAGGGAAGCGTCTTGCCTCTTTCCTTTGGTTTCTG
CCTGTACTCTGGTGTTTCCCGTGTCACCTGCATCATAGGAACAGCAGAGAAATCTGACCCAGTGCTATTT
TTCTAGGTGCTACTATGGCAAACTCAAGTGGTCTGTTTCTGTTCCTGTAACGTTCGACTATCTCGCTAGC
TGTGAAGTACTGATTAGTGGAGTTCTGTGCAACAGCAGTGTAGGAGTATACACAAACACAAATATGTGTT
TCTATTTAAAACTGTGGACTTAGCATAAAAAGGGAGAATATATTTATTTTTTACAAAAGGGATAAAAATG
GGCCCCGTTCCTCACCCACCAGATTTAGCGAGAAAAAGCTTTCTATTCTGAAAGGTCACGGTGGCTTTGG
CATTACAAATCAGAACAACACACACTGACCATGATGGCTTGTGAACTAACTGCAAGGCACTCCGTCATGG
TAAGCGAGTAGGTCCCACCTCCTAGTGTGCCGCTCATTGCTTTACACAGTAGAATCTTATTTGAGTGCTA
ATTGTTGTCTTTGCTGCTTTACTGTGTTGTTATAGAAAATGTAAGCTGTACAGTGAATAAGTTATTGAAG
CATGTGTAAACACTGTTATATATCTTTTCTCCTAGATGGGGAATTTTGAATAAAATACCTTTGAAATTCT
G

SEQ ID NO:124

>NM_001007702.1 大鼠(Rattus norvegicus),类似于 RIKEN cDNA 3110043O21 (RGD1359108), mRNA

[1124]

CGTTTGTAGTGTCAGCCATCCCAATTGCCTGTTCCTTCTCTGTGGGAGTGGTGTCTAGACAGTCCAGGCA
GGGTATGCTAGGCAGGTGCGTTTTGGTTGCCTCAGATCGCAACTTGACTCCATAACGGTGACCAAAGACA
AAAGAAGGAAACCAGATTAAAAAGAACCGGACACAGACCCCTGCAGAATCTGGAGCGGCCGTGGTTGGGG
GCGGGGCTACGACGGGGCGGACTCGGGGGCGTGGGAGGGCGGGGCCGGGGCGGGGCCCGGAGCCGGCTGC
GGTTGCGGTCCCTGCGCCGGCGGTGAAGGCGCAGCGGCGGCGAGTGGCTATTGCAAGCGTTTGGATAATG
TGAGACCTGGGATGCAGGGATGTCGACTATCTGCCCCCCACCATCTCCTGCTGTTGCCAAGACAGAGATT
GCTTTAAGTGGTGAATCACCCTTGTTGGCGGCTACCTTTGCTTACTGGGATAATATTCTTGGTCCTAGAG
TAAGGCACATTTGGGCTCCAAAGACAGACCAAGTACTCCTCAGTGATGGAGAAATCACTTTTCTTGCCAA
CCACACTCTGAATGGAGAAATTCTTCGGAATGCGGAGAGTGGGGCAATAGATGTAAAGTTTTTTGTCTTA
TCTGAAAAGGGCGTCATTATTGTTTCATTAATCTTCGACGGGAACTGGAACGGAGATCGGAGCACTTACG
GACTATCAATTATACTGCCGCAGACGGAGCTGAGTTTCTACCTCCCACTGCACAGAGTGTGTGTTGACAG
GCTAACGCACATCATTCGAAAAGGAAGGATATGGATGCACAAGGAAAGACAAGAAAATGTCCAGAAAATT
GTCTTGGAAGGCACCGAGAGGATGGAAGATCAGGGTCAGAGTATCATCCCTATGCTTACTGGGGAGGTCA
TCCCTGTGATGGAGCTGCTTGCGTCTATGAGATCACACAGTGTTCCTGAAGACCTCGATATAGCTGATAC
AGTACTCAATGATGATGACATTGGTGACAGCTGTCATGAAGGCTTTCTTCTCAATGCCATCAGCTCACAT
CTGCAGACCTGCGGCTGTTCTGTGGTGGTAGGCAGCAGTGCAGAGAAAGTAAATAAGATAGTAAGAACAC
TGTGCCTTTTTCTGACACCAGCAGAGAGGAAGTGCTCCAGGCTGTGTGAAGCCGAATCGTCCTTTAAATA
CGAATCTGGACTCTTTGTACAAGGCTTGCTAAAGGATGCGACTGGCAGTTTTGTACTACCTTTCCGGCAA
GTTATGTATGCCCCTTATCCCACCACACACATCGATGTGGATGTCAACACTGTCAAGCAGATGCCACCGT
GTCATGAACATATTTATAATCAACGCAGATACATGAGGTCAGAGCTGACAGCCTTCTGGAGGGCAACTTC
AGAAGAGGACATGGCTCAGGACACCATCATCTACACAGATGAGAGCTTCACTCCTGATTTGAATATTTTC
CAAGATGTCTTACACAGAGACACTCTAGTGAAAGCCTTTCTGGATCAGGTCTTCCATTTGAAGCCTGGCC
TGTCTCTCAGGAGTACTTTCCTTGCACAGTTCCTCCTCATTCTTCACAGAAAAGCCTTGACACTAATCAA
GTACATAGAGGATGACACGCAGAAGGGGAAAAAGCCCTTTAAGTCTCTTCGGAACCTGAAGATAGATCTT
GATTTAACAGCAGAGGGCGACCTTAACATAATAATGGCTCTAGCTGAGAAAATTAAGCCAGGCCTACACT
CTTTCATCTTCGGGAGACCTTTCTACACTAGTGTCCAAGAACGTGATGTTCTAATGACTTTTTAAACATG
TGGTTTGCTCCGTGTGTCTCATGACAGTCACACTTGCTGTTACAGTGTCTCAGCGCTTTGGACACATCCT
TCCTCCAGGGTCCTGCCGCAGGACACGTTACACTACACTTGTCAGTAGAGGTCTGTACCAGATGTCAGGT
ACATCGTTGTAGTGAATGTCTCTTTTCCTAGACTAGATGTACCCTCGTAGGGACTTATGTTTACAACCCT
CCTAAGTACTAGTGCTGTCTTGTAAGGATACGAATGAAGGGATGTAAACTTCACCACAACTGCTGGTTGG
TTTTGTTGTTTTTGTTTTTTGAAACTTATAATTCATGGTTTACATGCATCACACTGAAACCCTAGTTAGC
TTTTTACAGGTAAGCTGTGAGTTGACTGCCTGTCCCTGTGTTCTCTGGCCTGTACGATCTGTGGCGTGTA
GGATCACTTTTGCAACAACTAAAAACTAAAGCACTTTGTTTGCAGTTCTACAGAAAGCAACTTAGTCTGT
CTGCAGATTCGTTTTTGAAAGAAGACATGAGAAAGCGGAGTTTTAGGTGAAGTCAGTTGTTGGATCTTCC
TTTATAGACTTAGTCCTTTAGATGTGGTCTGTATAGACATGCCCAACCATCATGCATGGGCACTGAATAT
CGTGAACTGTGGTATGCTTTTTGTTGGTTTATTGTACTTCTGTCAAAGAAAGTGGCATTGGTTTTTATAA
TTGTTGCCAAGTTTTAAGGTTAATTTTCATTATTTTTGAGCCAAATTAAAATGTGCACCTCCTGTGCCTT
TCCCAATCTTGGAAAATATAATTTCTTGGCAGAAGGTCAGATTTCAGGGCCCAGTCACTTTCGTCTGACT
TCCCTTTGCACAGTCCGCCATGGGCCTGGCTTAGAAGTTCTTGTAAACTATGCCAGAGAGTACATTCGCT
GATAAAATCTTCTTTGCAGAGCAGGAGAGCTTCTTGCCTCTTTCCTTTCATTTCTGCCTGGACTTTGGTG
TTCTCCACGTTCCCTGCATCCTAAGGACAGCAGGAGAACTCTGACCCCAGTGCTATTTCTCTAGGTGCTA
TTGTGGCAAACTCAAGCGGTCCGTCTCTGTCCCTGTAACGTTCGTACCTTGCTGGCTGTGAAGTACTGAC
TGGTAAAGCTCCGTGCTACAGCAGTGTAGGGTATACACAAACACAAGTAAGTGTTTTATTTAAAACTGTG
GACTTAGCATAAAAAGGGAGACTATATTTATTTTTTACAAAAGGGATAAAAATGGAACCCTTTCCTCACC
CACCAGATTTAGTCAGAAAAAAACATTCTATTCTGAAAGGTCACAGTGGTTTTGACATGACACATCAGAA
CAACGCACACTGTCCATGATGGCTTATGAACTCCAAGTCACTCCATCATGGTAAATGGGTAGATCCCTCC
TTCTAGTGTGCCACACCATTGCTTCCCACAGTAGAATCTTATTTAAGTGCTAAGTGTTGTCTCTGCTGGT
TTACTCTGTTGTTTTAGAGAATGTAAGTTGTATAGTGAATAAGTTATTGAAGCATGTGTAAACACTGTTA
TACATCTTTTCTCCTAGATGGGGAATTTGGAATAAAATACCTTTAAAATTCAAAAAAAAAAAAAAAAAAA
AAAAA

SEQ ID NO:125

> SEQ ID NO:121 SEQ ID NO:1 的反向互补序列

TTTTTTTTTTTTTTTTTTAAAATTTCAAATATATTTTATTCAAAATTCTCCATTTAGGAGAAAAGATATATAACAAT
GTTTACACATGCTTTAATAACTTATTTCACTGTACAACTTACATTCTGTATAACAGTACAATAAACCAGCCAAAGAA
AATAACCAGTTAGCACTTAAATAAGAATCTACCATGTAAAAAACACAGTATGGGACACTACAAGGTAGTATTTATAT
ATTTTTTAAATGACTGAGCTACAGTACAACAGTCATCTAGTTCAGTGGTTGTCTAAAACATCAAGCTGTCCACATCT
TTCTGATTCATGATGGGAAAGCTATTATGACCTTTCACATTCGAACATGTCATTTTGTTGTGTAAATTTGGTGGGTG
GGGGGCAGAAGGGCTCTATTACCTTTATCCCTTTCTTATAAATATATTTTCCCTTTTATATTACTTCCAGAATTTTA
AATAAAATATTTATTTGTGTTTGTGTAACTACAATGTCAACATTGTGTTAGAATTATATTAATCAGTTATTTTATAG

[1125]

CCAGCAAAATGGTCCAAACGCATTAAGAAAACAAAAGATAACTTAGCTCTGCCAAAGTAGCACCTAGGAAAACAGCA
CTTCAGTAAAATTTCTCCTGCTATTCCTTATAGCCTATTGCAGGCAAACAGCAACAACTTCAAAAACATAGGCAGAA
ATGCAAGAAGGAAGTACTCAAGGAACTATTTTCTGTTTCTGAAAAAAAGATCTTATTAGTTAGTATATTCTCTAAGG
CATTTTATAATCTCAGAGTTGCAATGATTGCCAAAGCAGGTACATGGTAAGACTTAGCAAGAAGAAAACAAGATGAA
AATGACTGACCAGTGAAATCTGAACTTGTTCCAAGTAATTAGATTTTCTAAGGAGAAAAAAGGCACAGGAGGTGCAC
ATTTCAATTTGGCTCAAAAATAATGAAAATTAATTTAAAAAGTTGGCAACAATTACTAAAACATAAAATACAATTTC
CTTTTACAGAGCTCAACTACATAGAATATCAACAATAGCAAGAACAATAAAATAAACAAAACACCCTAACTGTGGTT
CACCAGGAACAAGGACTATAGAGAACTATTAGACATTTCTACAGACTGAATCCCAGGGACTAAATCCACAAAGTAGG
ATCTAATAATCAATTCCAATTAAACATCTGCTTGATCAATTTTCTTTCTAAAGCTCATCCTATGTTCAAGCTCACTT
TCTTTCTGTGGAATTGAAATCATTTAAAGGCTCTAGTTTTCAGTTGATTGCAGAAGTATTACAGTAATTCTACACAC
CAAAGAATGCCAAAAGATAGAAAGTCAACTATCTGTAAAAGCCAACTCAGATTTCACCTTGACACATGTAAACCATG
AATCATGTATTTCAAAAAAACAGTAGTTGTGGTCAAGTTTACATCCTATTATTATATCTTTAAAAGATAGCAATAAT
ATTTATTATATTGTAAACATAGGTCTGTATCCCAAAAGCATAAATCTAGGAAAAGAGACACCCAATGTAATGATGCA
CCTGACATCCCCTCACAGGCTCTTGTGAGAACTGTAGTGTAACTTACTTAACTGCAATTGCTGAGAGCAGAATTCTG
GAGTATGATCCAGGGGAACGTTTCCCCACACCACTGAGCTACTTTACCAGCGATCATGATTGTGATGGAATAGGCTT
ATTAAGTTACACATTTAAAAAGTCATTAGAACATCTCGTTCTTGCACACTAGTGTAGAAAGGTCTTCCAAAGATAAA
AGAGTGTAGGCCTGGTTTAATTTTCTCAGCCAGAGCCATTATTATGTTAAGATCGCCCTCTGCTGTTAAATCAAGGT
CTATCTTCAGGTTCCGAAGAGATTTAAAGGGCTTTTTTCCCTTCTGCGTATCGTCTTCTATATATTTTATTAGTGTC
AAGGCTTTTCTGTGAAGGACAAGTAGAAACTGTGCAAGGAAAGTACTTCTGAGAGATAAGCCAGGTTTCAGCTGAAA
GACCTGATCCAGGAAGGCTTTCACTAGAGTGTCTCTGTGTAAGACATCTTGAAAAATATTCAAATCAGGAGTAAAGC
TTTCGTCAGTGTAGATGATCGTATCCTGAGCCATGTCTTCTTCTGAAGTGGCTCTCCAGAAGGCTGTCAGCTCGGAT
CTCATGTATCTACGCTGATTATAAATATGTTCATGACAGGGTGGCATCTGCTTCACAGTATTGACATCCACATCTAT
GTGTGTGGTGGGATATGGAGCATACATGACTTGCCGGAAAGGCAGCACAAAGCTTCCAGTTGAATCCTTTAGCAGGC
CTTGTACAAAGAGCCCTGACTCATATTTAAATGATGATTCTGCTTCACATAACCTGGAGCATTTTCTCTCTGCTGGA
GTCAGAAAAAGGCATAATGTTCTGACTATCTTATTTACTTTCTCTGCACTGCTACCTACTACAACGGAACAGCCACA
GGTTTGCAAGTGTGAGCTGATGGCATTGAGAAGAAAGCCTTCATGACAGCTGTCACCAATATCATCATCATTGAGTA
CTGTATCAGCTATATCTATTTCTTCAGGAACACTGTGTGATTTCATAGATGAAAGCAGTTCCATTACAGGAATCACT
TCTCCAGTAAGCATTGGAATAATACTCTGACCCTGATCTTCCATTCTCTCTGTGCCTTCTAAGATAATCTTCTGGAC
ATTTTCTTGTCTTTCCTTATGCATCCATATTCTTCCTTTCCGGATTATATGTGTTAATCTATCAACACACACTCTAT
GAAGTGGGAGGTAGAAACTAAGTTCTGTCTGTGGAAGTATAATTGATAGTCCATATGTGCTGCGATCCCCATTCCAG
TTTCCATCAAAGATTAATGAAACAATAATCACTCCCTTTTCAGACAAGACAAAAAACTTTACATCTATAGCACCACT
CTCTGCATTTCGAAGGATTTCTCCATTTAGAGTGTGGTTGGCAAGAAAAGTTATTTCTCCATCACTGAGAAGTACCT
GTTCTGTCTTTGGAGCCCAAATGTGCCTTACTCTAGGACCAAGAATATTGTCCCAGTAAGCAAAAGTAGCTGCTAAT
AAAGGTGATTTGCCACTTAAAGCAATCTCTGTCTTGGCAACAGCTGGAGATGGCGGTGGGCAAAGAGTCGACATCAC
TGCATTCCAACTGTCACATTATCCAAATGCTCCGGAGATATCTCCTAAACCCACACCTGCTCTTGCTAGACCCCGCC
CCCAAAAGAGAAGCAACCGGGCAGCAGGGACGGCTGACACACCAAGCGTCATCTTTTACGTGGGCGGAACTTGTCGC
TGTTTGACGCACCTCTCTTTCCTAGCGGGACACCGTAGGTTACGT

SEQ ID NO:126

> SEQ ID NO:122 SEQ ID NO:2 的反向互补序列

AAAATTTCAAATATATTTTATTCAAAATTCTCCATCTAGGAGAAAAGATATATAACAATGTTTACACATGCTTTAAT
AACTTATTTCACTGTACAACTTACATTCTATATAACAGTACAATAAACCAGCCAAAGAAAATAACCAGGTAGCACTT
AAATAAGAATCTATCATGTAAAAAACATAGTATGGGACACTACGAGATAGTATTTATATATTTTTTAAATGACTGAG
CTACAGTACAACAGTCATCTAGTTCAGGGGTTGTCTGAAAAATCAAGCTGCCACATCTTTCTGATTAATGATGGGAA
AGCTATTATGACCTTTCACATTAGAACATATCATTTTGTTGTGTAAATTTGGTGGCTGGGGGACAGAAGAGCTCTAT
TACCTTTATTCCTTTCTTATAAATATATTTTCCCTTTTATGTTACTTCCAGTTTTAAATAAAATATTTGTGTTTGTG
TACACTACAATGTCAATATTGTGTTAGAATTATATTAATCAGTTATTTTACAGCCAGCAAAATGGTCCAAACGCATT
AAGGAAACAAAAGAAACAGACCACTTAGCTCTGCCAAAGTAGCACCTAGAAAAATAGCACTTCAGTAAAATTTCTCC
TGCTATTCCTTATAGCCTATTGCAGGCAAATGGCAACAACTTCAAAAACATGGGCAGAAATGTAAGAAGGAAGTACT
CAAGGAACTATTTTCTGTTCCTGAAAAAAAAGATCTTATTAGTTAATATATTCTCTAATGCATTTTATAATCTCAGA
GTTACAATGATTGCCAAAGCAGGTACATGGTAAGACTCTGCAAGAAGAAAACAAGATGAAAACGACTGACCAGTGAA
ATCTGAACTTCTTCCAAGTAATTCGATTTTCCAAGGAAAAAAAAGGCACAGGAGGTGCACATTTCAATTTGGCTCAA
AAATAATGAAAATTAATTTAAAAAGTTGGCAACAACTACTAAAACATAAAATACAATTTCCTTTTACAGAGCTCAAC
TATATAGAACATCAACAATAGCAACAACAATAAAACAAACAAAACATCCTAACTGTGGTTCACCATGAGCAGGGACT
ATAGAGAACTATCAGACATTTCTACAGACTGAATCCCGAGGGCTAAATCCACAAAGCAGGATCTAATAATCAATTCC
AATTAAACATCTGCTTGATCAATTCTCTTTCTAAAGCTTATCCTATGTTGAAGCTCACTTTCTTTCTGTGGAACTGA
AATCATTTAAAGGCTCTAATTTTCAGTTGATTGCAGAAGTATTACAGTAATTCCACACACCAAAGAATGCCAAAAGA
TAGAAAGTCAACTAATATCTGTAAAAGCGAACTCGGATTTCACCTTAATACATGTAAACCATGAATCATATATTTCA
AAAAAACAGCAGTTGTGGTCAAGTTTATATCCTATTATTATATATTTAAAAGACAGCAATAATACCTATTATATTGT
AAACATAGGTCTGTATCCCAAAAGCATAAATCTAGGAAAAGAGACATCCAATGTAATGATGCACCTGACATCCCCTC
ACAGACTCTTGTGAGAACTGTAGTGTAACTAACTTAACTGCAACTGCCGAGCAGAGCTCTGGAGTATGATCCAGGGG
AATGTTTCCCCACACCACCGAGCTACTTTAGCAGCGATCACGATTGTGATGGAATAGGCTTATTAAGTTACACATTT
AAAAAGTCATTAGAACATCTCGTTCTTGTACACTAGTGTAGAAAGGTCTTCCAAAGATAAAAGAGTGTAGGCCTGGT
TTAATTTTCTCAGCCAGAGCCATTATTATGTTAAGATCGCCCTCTGCTGTTAAATCAAGGTCTATCTTCAGGTTCCG
AAGAGATTTAAAGGGCTTTTTTCCCTTCTGCGTATCATCTTCTATATATTTTATTAGTGTCAAGGCTTTTCTGTGAA
GGACAAGTAAAAACTGTGCAAGGAAAGTACTCCTGAGAGATAAGCCAGGTTTCAGCTGAAAGACCTGATCCAGGAAG

GCTTTCACTAGAGTGTCTCTGTGTAAGACATCTTGAAAAATATTCAAATCAGGAGTAAAGCTTTCGTCAGTGTAGAT
GATCGTATCCTGAGCCATGTCTTCTTCTGAAGTGGCTCTCCAGAAGGCTGTCAGCTCGGATCTCATGTATCTACGCT
GATTATAAATATGTTCATGACAGGGTGGCATCTGCTTCACAGTATTGACATCCACATCTATGTGTGTGGTGGGATAT
GGAGCATACATGACTTGCCGGAAAGGCAGCACAAAGCTTCCAGTTGAATCCTTTAGCAGGCCCTGTACAAAGAGCCC
TGACTCATATTTAAATGATGATTCTGCTTCACATAACCTGGAGCATTTTCTCTCTGCTGCAGTCAGAAAAAGGCATA
ATGTTCTGACTATCTTATTTACTTTCTCTGCACTGCTACCTACTACAACGGAACAGCCACAGGTTTGCAAGTGTGAG
CTGATGGCATTGAGAAGAAAGCCTTCATGACAACTGTCACCAATATCATCATCATTGAGTACTGTATCAGCTATATC
TATTTCTTCAGGAACACTGTGTGATTTCATAGATGAAAGCAGTTCCATTACAGGAATCACTTCTCCAGTAAGCATTG
GAATAATACTCTGACCCTGATCTTCCATTCTCTCTGTGCCTTCTAAGATAATCTTCTGGACATTTTCTTGTCTTTCC
TTATGCATCCATATTCTTCCTTTCCGGATTATATGTGTTAATCTATCAACACACACTCTATGAAGTGGGAGGTAGAA
ACTAAGTTCTGTCTGTGGAAGTATAATTGATAGTCCGTATGTGCTGCGATCCCCATTCCAGTTTCCATCAAAGATTA
ATGAAACAATAATCACTCCCTTTTCAGACAAGACAAAAAACTTTACATCTATAGCACCACTCTCTGCATTTCGAAGG
ATTTCTCCATTTAGAGTGTGGTTGGCAAGAAAAGTTATTTCTCCGTCACTGAGAAGTACCTGTTCTGTCTTTGGAGC
CCAAATGTGCCTTACTCTAGGACCAAGAATATTGTCCCAGTAAGCAAAAGTAGCTGCTAATAAAGGTGATTCACCAC
TTAAAGCAATCTCTGTCTTGGCAACAGCTGGAGATGGCGGTGGGCAAAGAGTCGACATCACTGCATTCCAACTGTCA
CATTATCCAAATGCTCCTGAGATATCTCCTAAACCCACACCTGCTCTAGCCAGGCCCCGCCCCCAGAGAGAAGGAAC
CGGGCAGCAGGGACGGCTGACGCACCAAGCGTCATCTTTTACGTGGGCGGAACTTGTCGCTGTTTGACGCGCCTCTC
TTTCCTAGCGGGACACCGTAGGTTACGT

SEQ ID NO:127

> SEQ ID NO:123 SEQ ID NO:3 的反向互补序列

CAGAATTTCAAAGGTATTTTATTCAAAATTCCCCATCTAGGAGAAAAGATATATAACAGTGTTTACACATGCTTCAA
TAACTTATTCACTGTACAGCTTACATTTTCTATAACAACACAGTAAAGCAGCAAAGACAACAATTAGCACTCAAATA
AGATTCTACTGTGTAAAGCAATGAGCGGCACACTAGGAGGTGGGACCTACTCGCTTACCATGACGGAGTGCCTTGCA
GTTAGTTCACAAGCCATCATGGTCAGTGTGTGTTGTTCTGATTTGTAATGCCAAAGCCACCGTGACCTTTCAGAATA
GAAAGCTTTTTCTCGCTAAATCTGGTGGGTGAGGAACGGGGCCCATTTTTATCCCTTTTGTAAAAAATAAATATATT
CTCCCTTTTTATGCTAAGTCCACAGTTTTAAATAGAAACACATATTTGTGTTTGTGTATACTCCTACACTGCTGTTG
CACAGAACTCCACTAATCAGTACTTCACAGCTAGCGAGATAGTCGAACGTTACAGGAACAGAAACAGACCACTTGAG
TTTGCCATAGTAGCACCTAGAAAAATAGCACTGGGTCAGATTTCTCTGCTGTTCCTATGATGCAGGTGACACGGGAA
ACACCAGAGTACAGGCAGAAACCAAAGGAAAGAGGCAAGACGCTTCCCTGCTCTTCAAAGATCTTATCAGCGAATGT
ACTCTGACATACTTAACAAGGACTTCTAATCTAAGCCAGGCACACGGCAGCCGTGCAAAGGGTGGTCAGATGAAAGT
GACTGGGCCCTGAAATCTGACCCTCTGCCAAGAAATTATATTTTCCAAGATTGGGAAAGGCACAGGAGGTGCGCATT
TTAATTCGGCTCAAAAATAATGAAAATTAACCTTAAAAGTTGGCAACAATTATAAAAACAAATGCCACTTTCTTTGA
CAGAACTACAATAAACCAACAAAAAACTTGTCACAGGTCACGATACTCAGGGCCCATGCATGACCCCTCGGCATTTC
TGTACAGACCGAATCTGAAGGACTACGTCTATAAAGGAAGGGCTAACAACTGATTTCACCTACAGCTCTGCCCTATC
AGTGTCTTCTCTCAAGAACTAATCTGTAGTCAGACTAAGTTCCTTTCTGAGGAACTGCAAACTTTAAATGCTCTAAT
GTTTAGCTGTGGCAGAAGTAATCGGACATACCACAGAACACAGAGGGGCAGGAAGTCAACTCACACTTTATCTGTAA
AGCTAACGAGGGTTTCAGTGTGGTACATGTAAACCAGGAATCATAAGTTTCCAAGAAAACAAACAAAACAAAACAAA
AAACCAATCAGCAGTTATGGTCAAGTGTATATTCCATCATTCATATCTTTAAAAGACAGCAATAGTACTTAGGATAA
TTGTAAACATAAGGCCCTAAAAGAGTACATCTATTCTAGGAAGAGTCATTCACTGTAACGATGCACCTGACATCAGG
TACAGACTTCTACTGACAAGTGTAGTGTACCCTGTCCTGCAGCAGGACTCGGGAGGAAGGATGTGCCCACACACTGC
TGAGACACTGTAACAGCAGGTGTGACTGTGAAGAGACAGACACAGCAAACCACACGGTCAAAAGGTCATTAGAACAT
CACGTTCTTGTACACTAGTGTAGAAAGGTCTCCCAAAGATGAAAGAGTGTAGGCCTGGCTTAATTTTCTCAGCTAGA
GCCATTATTATGTTAAGATCGCCCTCTGCTGTTAAATCAAGATCTATCTTCAGGTTCCGAAGAGACTTAAAGGGCTT
TTTCCCCTTCTGCGTATCATCCTCGATGTACTTGATTAGTGTCAAGGCTTTTCTGTGAAGAATGAGGAGGAACTGTG
CAAGGAAAGTACTCCTGAGAGACAGGCCAGGCTTCAAATGGAAGACCTGATCCAGGAAGGCTTTCACTAGAGTGTCT
CTGTGTAAGACATCTTGGAAAATATTCAAATCAGGAGTGAAGCTCTCATCTGTGTAGATGATGGTGTCCTGCGCCAT
GTCCTCTTCTGAAGTTGCCCTCCAGAAGGCTGTCAGCTCTGACCTCATGTATCTGCGTTGATTATAAATATGTTCAT
GACACGGTGGCATCTGCTTGACAGTGTTGACATCCACATCAATGTGCGTGGTGGGGTACGGGGCATACATAACTTGC
CGGAAGGGTAGGACAAAACTGCCTGTTGCATCCTTTAGCAAGCCTTGCACAAAGAGTCCCGATTCGTACTTAAAGGA
CGATTCTGCTTCACACAGCCTGGAGCATTTCCTCTCTGCTGGTGTCAGAAAAAGGCACAGCGTTCTTACTATCTTAT
TTACTTTCTCTGCACTGCTGCCAACTACAACGGAACAGCCACAGGTCTGCAGGTGTGAGCTGATGGCATTGAGAAGA
AAGCCTTCGTGACAGCTGTCACCAATGTCATCATCATTGAGCACTGTATCAGCTATATCAATGTCTTCAGGAACACT
GTGGGATTTCATAGATGCAAGCAGCTCCATTACAGGAATGACTTCCCCAGTAAGCATGGGAATGATACTCTGACCCT
GATCTTCCATCCTCTCTGTGCCTTCCAAGACAATTTTCTGGACATTTTCTTGTCTTTCCTTATGCATCCATATTCTT
CCTTTTCGAATAATGTGTGTTAGCCTGTCAACACACACTCTGTGAAGTGGGAGGTAGAAGCTCAGCTCTGTCTGCGG
CAGTATAATTGATAGTCCATAAGTGCTCCGATCTCCATTCCAGTTTCCGTCGAAGATTAATGAAACAATAATTACCC
CTTTTTCAGATAAGACAAAAAATTTTACATCTATAGCCCCACTCTCTGCATTTCGAAGAATTTCTCCATTTAGAGTG
TGGTTGGCAAGAAAAGTTATTTCTCCATCACTGAGAAGCACTTGGTCTGTCTTTGGAGCCCAAATATGCCTTACTCT
AGGACCAAGAATATTATCCCAGTAAGCAAAGGTAGCCGCCAACAAGGGTGATTCACCACTTAAAGCAATCTCTGTCT
TGGCAACAGCAGGAGATGGTGGGGGGCAGATAGTCGACATCCCTGCATCCCAGGTCTCACTGCATTCCAGGTCTCAC
ATTATCCGAACGCTTGCAATAGCCACTCGCCGCTGCTGCGCCTTCACCGCCGGCGCAGGGACCGCAACCGC

SEQ ID NO:128

> SEQ ID NO:124 SEQ ID NO:4 的反向互补序列

[1126]

TTTTTTTTTTTTTTTTTTTTTTTTGAATTTTAAAGGTATTTTATTCCAAATTCCCCATCTAGGAGAAAAGATGTATA
ACAGTGTTTACACATGCTTCAATAACTTATTCACTATACAACTTACATTCTCTAAAACAACAGAGTAAACCAGCAGA
GACAACACTTAGCACTTAAATAAGATTCTACTGTGGGAAGCAATGGTGTGGCACACTAGAAGGAGGGATCTACCCAT
TTACCATGATGGAGTGACTTGGAGTTCATAAGCCATCATGGACAGTGTGCGTTGTTCTGATGTGTCATGTCAAAACC
ACTGTGACCTTTCAGAATAGAATGTTTTTTTCTGACTAAATCTGGTGGGTGAGGAAAGGGTTCCATTTTTATCCCTT
TTGTAAAAAATAAATATAGTCTCCCTTTTTATGCTAAGTCCACAGTTTTAAATAAAACACTTACTTGTGTTTGTGTA
TACCCTACACTGCTGTAGCACGGAGCTTTACCAGTCAGTACTTCACAGCCAGCAAGGTACGAACGTTACAGGGACAG
AGACGGACCGCTTGAGTTTGCCACAATAGCACCTAGAGAAATAGCACTGGGGTCAGAGTTCTCCTGCTGTCCTTAGG
ATGCAGGGAACGTGGAGAACACCAAAGTCCAGGCAGAAATGAAAGGAAAGAGGCAAGAAGCTCTCCTGCTCTGCAAA
GAAGATTTTATCAGCGAATGTACTCTCTGGCATAGTTTACAAGAACTTCTAAGCCAGGCCCATGGCGGACTGTGCAA
AGGGAAGTCAGACGAAAGTGACTGGGCCCTGAAATCTGACCTTCTGCCAAGAAATTATATTTTCCAAGATTGGGAAA
GGCACAGGAGGTGCACATTTTAATTTGGCTCAAAAATAATGAAAATTAACCTTAAAACTTGGCAACAATTATAAAAA
CCAATGCCACTTTCTTTGACAGAAGTACAATAAACCAACAAAAAGCATACCACAGTTCACGATATTCAGTGCCCATG
CATGATGGTTGGGCATGTCTATACAGACCACATCTAAAGGACTAAGTCTATAAAGGAAGATCCAACAACTGACTTCA
CCTAAAACTCCGCTTTCTCATGTCTTCTTTCAAAAACGAATCTGCAGACAGACTAAGTTGCTTTCTGTAGAACTGCA
AACAAAGTGCTTTAGTTTTTAGTTGTTGCAAAAGTGATCCTACACGCCACAGATCGTACAGGCCAGAGAACACAGGG
ACAGGCAGTCAACTCACAGCTTACCTGTAAAAAGCTAACTAGGGTTTCAGTGTGATGCATGTAAACCATGAATTATA
AGTTTCAAAAAACAAAAACAACAAAACCAACCAGCAGTTGTGGTGAAGTTTACATCCCTTCATTCGTATCCTTACAA
GACAGCACTAGTACTTAGGAGGGTTGTAAACATAAGTCCCTACGAGGGTACATCTAGTCTAGGAAAAGAGACATTCA
CTACAACGATGTACCTGACATCTGGTACAGACCTCTACTGACAAGTGTAGTGTAACGTGTCCTGCGGCAGGACCCTG
GAGGAAGGATGTGTCCAAAGCGCTGAGACACTGTAACAGCAAGTGTGACTGTCATGAGACACACGGAGCAAACCACA
TGTTTAAAAAGTCATTAGAACATCACGTTCTTGGACACTAGTGTAGAAAGGTCTCCCGAAGATGAAAGAGTGTAGGC
CTGGCTTAATTTTCTCAGCTAGAGCCATTATTATGTTAAGGTCGCCCTCTGCTGTTAAATCAAGATCTATCTTCAGG
TTCCGAAGAGACTTAAAGGGCTTTTTCCCCTTCTGCGTGTCATCCTCTATGTACTTGATTAGTGTCAAGGCTTTTCT
GTGAAGAATGAGGAGGAACTGTGCAAGGAAAGTACTCCTGAGAGACAGGCCAGGCTTCAAATGGAAGACCTGATCCA
GAAAGGCTTTCACTAGAGTGTCTCTGTGTAAGACATCTTGGAAAATATTCAAATCAGGAGTGAAGCTCTCATCTGTG
TAGATGATGGTGTCCTGAGCCATGTCCTCTTCTGAAGTTGCCCTCCAGAAGGCTGTCAGCTCTGACCTCATGTATCT
GCGTTGATTATAAATATGTTCATGACACGGTGGCATCTGCTTGACAGTGTTGACATCCACATCGATGTGTGTGGTGG
GATAAGGGGCATACATAACTTGCCGGAAAGGTAGTACAAAACTGCCAGTCGCATCCTTTAGCAAGCCTTGTACAAAG
AGTCCAGATTCGTATTTAAAGGACGATTCGGCTTCACACAGCCTGGAGCACTTCCTCTCTGCTGGTGTCAGAAAAAG
GCACAGTGTTCTTACTATCTTATTTACTTTCTCTGCACTGCTGCCTACCACCACAGAACAGCCGCAGGTCTGCAGAT
GTGAGCTGATGGCATTGAGAAGAAAGCCTTCATGACAGCTGTCACCAATGTCATCATCATTGAGTACTGTATCAGCT
ATATCGAGGTCTTCAGGAACACTGTGTGATCTCATAGACGCAAGCAGCTCCATCACAGGGATGACCTCCCCAGTAAG
[1127] CATAGGGATGATACTCTGACCCTGATCTTCCATCCTCTCGGTGCCTTCCAAGACAATTTTCTGGACATTTTCTTGTC
TTTCCTTGTGCATCCATATCCTTCCTTTTCGAATGATGTGCGTTAGCCTGTCAACACACACTCTGTGCAGTGGGAGG
TAGAAACTCAGCTCCGTCTGCGGCAGTATAATTGATAGTCCGTAAGTGCTCCGATCTCCGTTCCAGTTCCCGTCGAA
GATTAATGAAACAATAATGACGCCCTTTTCAGATAAGACAAAAAACTTTACATCTATTGCCCCACTCTCCGCATTCC
GAAGAATTTCTCCATTCAGAGTGTGGTTGGCAAGAAAAGTGATTTCTCCATCACTGAGGAGTACTTGGTCTGTCTTT
GGAGCCCAAATGTGCCTTACTCTAGGACCAAGAATATTATCCCAGTAAGCAAAGGTAGCCGCCAACAAGGGTGATTC
ACCACTTAAAGCAATCTCTGTCTTGGCAACAGCAGGAGATGGTGGGGGGCAGATAGTCGACATCCCTGCATCCCAGG
TCTCACATTATCCAAACGCTTGCAATAGCCACTCGCCGCCGCTGCGCCTTCACCGCCGGCGCAGGGACCGCAACCGC
AGCCGGCTCCGGGCCCCGCCCCGGCCCCGCCCTCCCACGCCCCCGAGTCCGCCCCGTCGTAGCCCCGCCCCCAACCA
CGGCCGCTCCAGATTCTGCAGGGGTCTGTGTCCGGTTCTTTTTAATCTGGTTTCCTTCTTTTGTCTTTGGTCACCGT
TATGGAGTCAAGTTGCGATCTGAGGCAACCAAAACGCACCTGCCTAGCATACCCTGCCTGGACTGTCTAGACACCAC
TCCCACAGAGAAGGAACAGGCAATTGGGATGGCTGACACTACAAACG

SEQ ID NO:129
>NM_145005.6 智人 (Homo sapiens) C9orf72-SMCR8 复合物亚基 (C9orf72), 转录本变体 1, mRNA
ACGTAACCTACGGTGTCCCGCTAGGAAAGAGAGGTGCGTCAAACAGCGACAAGTTCCGCCCACGTAAAAG
ATGACGCTTGATATCTCCGGAGCATTTGGATAATGTGACAGTTGGAATGCAGTGATGTCGACTCTTTGCC
CACCGCCATCTCCAGCTGTTGCCAAGACAGAGATTGCTTTAAGTGGCAAATCACCTTTATTAGCAGCTAC
TTTTGCTTACTGGGACAATATTCTTGGTCCTAGAGTAAGGCACATTTGGGCTCCAAAGACAGAACAGGTA
CTTCTCAGTGATGGAGAAATAACTTTTCTTGCCAACCACACTCTAAATGGAGAAATCCTTCGAAATGCAG
AGAGTGGTGCTATAGATGTAAAGTTTTTTGTCTTGTCTGAAAAGGGAGTGATTATTGTTTCATTAATCTT
TGATGGAAACTGGAATGGGGATCGCAGCACATATGGACTATCAATTATACTTCCACAGACAGAACTTAGT
TTCTACCTCCCACTTCATAGAGTGTGTGTTGATAGATTAACACATATAATCCGGAAAGGAAGAATATGGA
TGCATAAGGAAAGACAAGAAAATGTCCAGAAGATTATCTTAGAAGGCACAGAGAGAATGGAAGATCAGGG
TCAGAGTATTATTCCAATGCTTACTGGAGAAGTGATTCCTGTAATGGAACTGCTTTCATCTATGAAATCA
CACAGTGTTCCTGAAGAAATAGATATAGCTGATACAGTACTCAATGATGATGATATTGGTGACAGCTGTC
ATGAAGGCTTTCTTCTCAAGTAAGAATTTTTCTTTTCATAAAAGCTGGATGAAGCAGATACCATCTTATG
CTCACCTATGACAAGATTTGGAAGAAAGAAAATAACAGACTGTCTACTTAGATTGTTCTAGGGACATTAC
GTATTTGAACTGTTGCTTAAATTTGTGTTATTTTTCACTCATTATATTTCTATATATATTTGGTGTTATT
CCATTTGCTATTTAAAGAAACCGAGTTTCCATCCCAGACAAGAAATCATGGCCCCTTGCTTGATTCTGGT
TTCTTGTTTTACTTCTCATTAAAGCTAACAGAATCCTTTCATATTAAGTTGTACTGTAGATGAACTTAAG
TTATTTAGGCGTAGAACAAAATTATTCATATTTATACTGATCTTTTTCCATCCAGCAGTGGAGTTTAGTA

CTTAAGAGTTTGTGCCCTTAAACCAGACTCCCTGGATTAATGCTGTGTACCCGTGGGCAAGGTGCCTGAA
TTCTCTATACACCTATTTCCTCATCTGTAAAATGGCAATAATAGTAATAGTACCTAATGTGTAGGGTTGT
TATAAGCATTGAGTAAGATAAATAATATAAAGCACTTAGAACAGTGCCTGGAACATAAAAACACTTAATA
ATAGCTCATAGCTAACATTTCCTATTTACATTTCTTCTAGAAATAGCCAGTATTTGTTGAGTGCCTACAT
GTTAGTTCCTTTACTAGTTGCTTTACATGTATTATCTTATATTCTGTTTTAAAGTTTCTTCACAGTTACA
GATTTTCATGAAATTTTACTTTTAATAAAAGAGAAGTAAAAGTATAAAGTATTCACTTTTATGTTCACAG
TCTTTTCCTTTAGGCTCATGATGGAGTATCAGAGGCATGAGTGTGTTTAACCTAAGAGCCTTAATGGCTT
GAATCAGAAGCACTTTAGTCCTGTATCTGTTCAGTGTCAGCCTTTCATACATCATTTTAAATCCCATTTG
ACTTTAAGTAAGTCACTTAATCTCTCTACATGTCAATTTCTTCAGCTATAAAATGATGGTATTTCAATAA
ATAAATACATTAATTAAATGATATTATACTGACTAATTGGGCTGTTTTAAGGCTCAATAAGAAAATTTCT
GTGAAAGGTCTCTAGAAAATGTAGGTTCCTATACAAATAAAAGATAACATTGTGCTTATAAAAAAAA

SEQ ID NO:130

SEQ ID NO:129 SEQ ID NO:9 的反向互补序列

TTTTTTTTATAAGCACAATGTTATCTTTTATTTGTATAGGAACCTACATTTTCTAGAGACCTTTCACAGAAATTTTC
TTATTGAGCCTTAAAACAGCCCAATTAGTCAGTATAATATCATTTAATTAATGTATTTATTTATTGAAATACCATCA
TTTTATAGCTGAAGAAATTGACATGTAGAGAGATTAAGTGACTTACTTAAAGTCAAATGGGATTTAAAATGATGTAT
GAAAGGCTGACACTGAACAGATACAGGACTAAAGTGCTTCTGATTCAAGCCATTAAGGCTCTTAGGTTAAACACACT
CATGCCTCTGATACTCCATCATGAGCCTAAAGGAAAAGACTGTGAACATAAAAGTGAATACTTTATACTTTTACTTC
TCTTTTATTAAAAGTAAAATTTCATGAAAATCTGTAACTGTGAAGAAACTTTAAAACAGAATATAAGATAATACATG
TAAAGCAACTAGTAAAGGAACTAACATGTAGGCACTCAACAAATACTGGCTATTTCTAGAAGAAATGTAAATAGGAA
ATGTTAGCTATGAGCTATTATTAAGTGTTTTTATGTTCCAGGCACTGTTCTAAGTGCTTTATATTATTTATCTTACT
CAATGCTTATAACAACCCTACACATTAGGTACTATTACTATTATTGCCATTTTACAGATGAGGAAATAGGTGTATAG
AGAATTCAGGCACCTTGCCCACGGGTACACAGCATTAATCCAGGGAGTCTGGTTTAAGGGCACAAACTCTTAAGTAC
TAAACTCCACTGCTGGATGGAAAAAGATCAGTATAAATATGAATAATTTTGTTCTACGCCTAAATAACTTAAGTTCA
TCTACAGTACAACTTAATATGAAAGGATTCTGTTAGCTTTAATGAGAAGTAAAACAAGAAACCAGAATCAAGCAAGG
GGCCATGATTTCTTGTCTGGGATGGAAACTCGGTTTCTTTAAATAGCAAATGGAATAACACCAAATATATATAGAAA
TATAATGAGTGAAAAATAACACAAATTTAAGCAACAGTTCAAATACGTAATGTCCCTAGAACAATCTAAGTAGACAG
TCTGTTATTTTCTTTCTTCCAAATCTTGTCATAGGTGAGCATAAGATGGTATCTGCTTCATCCAGCTTTTATGAAAA
GAAAAATTCTTACTTGAGAAGAAAGCCTTCATGACAGCTGTCACCAATATCATCATCATTGAGTACTGTATCAGCTA
TATCTATTTCTTCAGGAACACTGTGTGATTTCATAGATGAAAGCAGTTCCATTACAGGAATCACTTCTCCAGTAAGC
ATTGGAATAATACTCTGACCCTGATCTTCCATTCTCTCTGTGCCTTCTAAGATAATCTTCTGGACATTTTCTTGTCT
[1128] TTCCTTATGCATCCATATTCTTCCTTTCCGGATTATATGTGTTAATCTATCAACACACACTCTATGAAGTGGGAGGT
AGAAACTAAGTTCTGTCTGTGGAAGTATAATTGATAGTCCATATGTGCTGCGATCCCCATTCCAGTTTCCATCAAAG
ATTAATGAAACAATAATCACTCCCTTTTCAGACAAGACAAAAAACTTTACATCTATAGCACCACTCTCTGCATTTCG
AAGGATTTCTCCATTTAGAGTGTGGTTGGCAAGAAAAGTTATTTCTCCATCACTGAGAAGTACCTGTTCTGTCTTTG
GAGCCCAAATGTGCCTTACTCTAGGACCAAGAATATTGTCCCAGTAAGCAAAAGTAGCTGCTAATAAAGGTGATTTG
CCACTTAAAGCAATCTCTGTCTTGGCAACAGCTGGAGATGGCGGTGGGCAAAGAGTCGACATCACTGCATTCCAACT
GTCACATTATCCAAATGCTCCGGAGATATCAAGCGTCATCTTTTACGTGGGCGGAACTTGTCGCTGTTTGACGCACC
TCTCTTTCCTAGCGGGACACCGTAGGTTACGT

SEQ ID NO:131

>NM_018325.5 智人 (Homo sapiens) C9orf72-SMCR8 复合物亚基 (C9orf72), 转录本变体 2, mRNA

GGTTGCGGTGCCTGCGCCCGCGGCGGCGGAGGCGCAGGCGGTGGCGAGTGGATATCTCCGGAGCATTTGG
ATAATGTGACAGTTGGAATGCAGTGATGTCGACTCTTTGCCCACCGCCATCTCCAGCTGTTGCCAAGACA
GAGATTGCTTTAAGTGGCAAATCACCTTTATTAGCAGCTACTTTTGCTTACTGGGACAATATTCTTGGTC
CTAGAGTAAGGCACATTTGGGCTCCAAAGACAGAACAGGTACTTCTCAGTGATGGAGAAATAACTTTTCT
TGCCAACCACACTCTAAATGGAGAAATCCTTCGAAATGCAGAGAGTGGTGCTATAGATGTAAAGTTTTTT
GTCTTGTCTGAAAAGGGAGTGATTATTGTTTCATTAATCTTTGATGGAAACTGGAATGGGGATCGCAGCA
CATATGGACTATCAATTATACTTCCACAGACAGAACTTAGTTTCTACCTCCCACTTCATAGAGTGTGTGT
TGATAGATTAACACATATAATCCGGAAAGGAAGAATATGGATGCATAAGGAAAGACAAGAAAATGTCCAG
AAGATTATCTTAGAAGGCACAGAGAGAATGGAAGATCAGGGTCAGAGTATTATTCCAATGCTTACTGGAG
AAGTGATTCCTGTAATGGAACTGCTTTCATCTATGAAATCACACAGTGTTCCTGAAGAAATAGATATAGC
TGATACAGTACTCAATGATGATGATATTGGTGACAGCTGTCATGAAGGCTTTCTTCTCAATGCCATCAGC
TCACACTTGCAAACCTGTGGCTGTTCCGTTGTAGTAGGTAGCAGTGCAGAGAAAGTAAATAAGATAGTCA
GAACATTATGCCTTTTTCTGACTCCAGCAGAGAGAAAATGCTCCAGGTTATGTGAAGCAGAATCATCATT
TAAATATGAGTCAGGGCTCTTTGTACAAGGCCTGCTAAAGGATTCAACTGGAAGCTTTGTGCTGCCTTTC
CGGCAAGTCATGTATGCTCCATATCCCACCACACACATAGATGTGGATGTCAATACTGTGAAGCAGATGC
CACCCTGTCATGAACATATTTATAATCAGCGTAGATACATGAGATCCGAGCTGACAGCCTTCTGGAGAGC
CACTTCAGAAGAAGACATGGCTCAGGATACGATCATCTACACTGACGAAAGCTTTACTCCTGATTTGAAT
ATTTTTCAAGATGTCTTACACAGAGACACTCTAGTGAAAGCCTTCCTGGATCAGGTCTTTCAGCTGAAAC
CTGGCTTATCTCTCAGAAGTACTTTCCTTGCACAGTTTCTACTTGTCCTTCACAGAAAAGCCTTGACACT
AATAAAATATATAGAAGACGATACGCAGAAGGGAAAAAAGCCCTTTAAATCTCTTCGGAACCTGAAGATA

GACCTTGATTTAACAGCAGAGGGCGATCTTAACATAATAATGGCTCTGGCTGAGAAAATTAAACCAGGCC
TACACTCTTTTATCTTTGGAAGACCTTTCTACACTAGTGTGCAAGAACGAGATGTTCTAATGACTTTTTA
AATGTGTAACTTAATAAGCCTATTCCATCACAATCATGATCGCTGGTAAAGTAGCTCAGTGGTGTGGGGA
AACGTTCCCCTGGATCATACTCCAGAATTCTGCTCTCAGCAATTGCAGTTAAGTAAGTTACACTACAGTT
CTCACAAGAGCCTGTGAGGGGATGTCAGGTGCATCATTACATTGGGTGTCTCTTTTCCTAGATTTATGCT
TTTGGGATACAGACCTATGTTTACAATATAATAAATATTATTGCTATCTTTTAAAGATATAATAATAGGA
TGTAAACTTGACCACAACTACTGTTTTTTTGAAATACATGATTCATGGTTTACATGTGTCAAGGTGAAAT
CTGAGTTGGCTTTTACAGATAGTTGACTTTCTATCTTTTGGCATTCTTTGGTGTGTAGAATTACTGTAAT
ACTTCTGCAATCAACTGAAAACTAGAGCCTTTAAATGATTTCAATTCCACAGAAAGAAAGTGAGCTTGAA
CATAGGATGAGCTTTAGAAAGAAAATTGATCAAGCAGATGTTTAATTGGAATTGATTATTAGATCCTACT
TTGTGGATTTAGTCCCTGGGATTCAGTCTGTAGAAATGTCTAATAGTTCTCTATAGTCCTTGTTCCTGGT
GAACCACAGTTAGGGTGTTTTGTTTATTTTATTGTTCTTGCTATTGTTGATATTCTATGTAGTTGAGCTC
TGTAAAAGGAAATTGTATTTTATGTTTTAGTAATTGTTGCCAACTTTTTAAATTAATTTTCATTATTTTT
GAGCCAAATTGAAATGTGCACCTCCTGTGCCTTTTTTCTCCTTAGAAAATCTAATTACTTGGAACAAGTT
CAGATTTCACTGGTCAGTCATTTTCATCTTGTTTTCTTCTTGCTAAGTCTTACCATGTACCTGCTTTGGC
AATCATTGCAACTCTGAGATTATAAAATGCCTTAGAGAATATACTAACTAATAAGATCTTTTTTTCAGAA
ACAGAAAATAGTTCCTTGAGTACTTCCTTCTTGCATTTCTGCCTATGTTTTTGAAGTTGTTGCTGTTTGC
CTGCAATAGGCTATAAGGAATAGCAGGAGAAATTTTACTGAAGTGCTGTTTTCCTAGGTGCTACTTTGGC
AGAGCTAAGTTATCTTTTGTTTTCTTAATGCGTTTGGACCATTTTGCTGGCTATAAAATAACTGATTAAT
ATAATTCTAACACAATGTTGACATTGTAGTTACACAAACACAAATAAATATTTTATTTAAAATTCTGGAA
GTAATATAAAAGGGAAAATATATTTATAAGAAAGGGATAAAGGTAATAGAGCCCTTCTGCCCCCCACCCA
CCAAATTTACACAACAAAATGACATGTTCGAATGTGAAAGGTCATAATAGCTTTCCCATCATGAATCAGA
AAGATGTGGACAGCTTGATGTTTTAGACAACCACTGAACTAGATGACTGTTGTACTGTAGCTCAGTCATT
TAAAAAATATATAAATACTACCTTGTAGTGTCCCATACTGTGTTTTTTACATGGTAGATTCTTATTTAAG
TGCTAACTGGTTATTTTCTTTGGCTGGTTTATTGTACTGTTATACAGAATGTAAGTTGTACAGTGAAATA
AGTTATTAAAGCATGTGTAAACATTGTTATATATCTTTTCTCCTAAATGGAGAATTTTGAATAAAATATA
TTTGAAATTTT

SEQ ID NO:132

SEQ ID NO:131 SEQ ID NO:11 的反向互补序列

[1129]

AAAATTTCAAATATATTTTATTCAAAATTCTCCATTTAGGAGAAAAGATATATAACAATGTTTACACATGCTTTAAT
AACTTATTTCACTGTACAACTTACATTCTGTATAACAGTACAATAAACCAGCCAAAGAAAATAACCAGTTAGCACTT
AAATAAGAATCTACCATGTAAAAAACACAGTATGGGACACTACAAGGTAGTATTTATATATTTTTTAAATGACTGAG
CTACAGTACAACAGTCATCTAGTTCAGTGGTTGTCTAAAACATCAAGCTGTCCACATCTTTCTGATTCATGATGGGA
AAGCTATTATGACCTTTCACATTCGAACATGTCATTTTGTTGTGTAAATTTGGTGGGTGGGGGGCAGAAGGGCTCTA
TTACCTTTATCCCTTTCTTATAAATATATTTTCCCTTTTATATTACTTCCAGAATTTTAAATAAAATATTTATTTGT
GTTTGTGTAACTACAATGTCAACATTGTGTTAGAATTATATTAATCAGTTATTTTATAGCCAGCAAAATGGTCCAAA
CGCATTAAGAAAACAAAAGATAACTTAGCTCTGCCAAAGTAGCACCTAGGAAAACAGCACTTCAGTAAAATTTCTCC
TGCTATTCCTTATAGCCTATTGCAGGCAAACAGCAACAACTTCAAAAACATAGGCAGAAATGCAAGAAGGAAGTACT
CAAGGAACTATTTTCTGTTTCTGAAAAAAAGATCTTATTAGTTAGTATATTCTCTAAGGCATTTTATAATCTCAGAG
TTGCAATGATTGCCAAAGCAGGTACATGGTAAGACTTAGCAAGAAGAAAACAAGATGAAAATGACTGACCAGTGAAA
TCTGAACTTGTTCCAAGTAATTAGATTTTCTAAGGAGAAAAAAGGCACAGGAGGTGCACATTTCAATTTGGCTCAAA
AATAATGAAAATTAATTTAAAAAGTTGGCAACAATTACTAAAACATAAAATACAATTTCCTTTTACAGAGCTCAACT
ACATAGAATATCAACAATAGCAAGAACAATAAAATAAACAAAACACCCTAACTGTGGTTCACCAGGAACAAGGACTA
TAGAGAACTATTAGACATTTCTACAGACTGAATCCCAGGGACTAAATCCACAAAGTAGGATCTAATAATCAATTCCA
ATTAAACATCTGCTTGATCAATTTTCTTTCTAAAGCTCATCCTATGTTCAAGCTCACTTTCTTTCTGTGGAATTGAA
ATCATTTAAAGGCTCTAGTTTTCAGTTGATTGCAGAAGTATTACAGTAATTCTACACACCAAAGAATGCCAAAAGAT
AGAAAGTCAACTATCTGTAAAAGCCAACTCAGATTTCACCTTGACACATGTAAACCATGAATCATGTATTTCAAAAA
AACAGTAGTTGTGGTCAAGTTTACATCCTATTATTATATCTTTAAAAGATAGCAATAATATTTATTATATTGTAAAC
ATAGGTCTGTATCCCAAAAGCATAAATCTAGGAAAAGAGACACCCAATGTAATGATGCACCTGACATCCCCTCACAG
GCTCTTGTGAGAACTGTAGTGTAACTTACTTAACTGCAATTGCTGAGAGCAGAATTCTGGAGTATGATCCAGGGGAA
CGTTTCCCCACACCACTGAGCTACTTTACCAGCGATCATGATTGTGATGGAATAGGCTTATTAAGTTACACATTTAA
AAAGTCATTAGAACATCTCGTTCTTGCACACTAGTGTAGAAAGGTCTTCCAAAGATAAAAGAGTGTAGGCCTGGTTT
AATTTTCTCAGCCAGAGCCATTATTATGTTAAGATCGCCCTCTGCTGTTAAATCAAGGTCTATCTTCAGGTTCCGAA
GAGATTTAAAGGGCTTTTTTCCCTTCTGCGTATCGTCTTCTATATATTTTATTAGTGTCAAGGCTTTTCTGTGAAGG
ACAAGTAGAAACTGTGCAAGGAAAGTACTTCTGAGAGATAAGCCAGGTTTCAGCTGAAAGACCTGATCCAGGAAGGC
TTTCACTAGAGTGTCTCTGTGTAAGACATCTTGAAAAATATTCAAATCAGGAGTAAAGCTTTCGTCAGTGTAGATGA
TCGTATCCTGAGCCATGTCTTCTTCTGAAGTGGCTCTCCAGAAGGCTGTCAGCTCGGATCTCATGTATCTACGCTGA
TTATAAATATGTTCATGACAGGGTGGCATCTGCTTCACAGTATTGACATCCACATCTATGTGTGTGGTGGGATATGG
AGCATACATGACTTGCCGGAAAGGCAGCACAAAGCTTCCAGTTGAATCCTTTAGCAGGCCTTGTACAAAGAGCCCTG
ACTCATATTTAAATGATGATTCTGCTTCACATAACCTGGAGCATTTTCTCTCTGCTGGAGTCAGAAAAAGGCATAAT
GTTCTGACTATCTTATTTACTTTCTCTGCACTGCTACCTACTACAACGGAACAGCCACAGGTTTGCAAGTGTGAGCT
GATGGCATTGAGAAGAAAGCCTTCATGACAGCTGTCACCAATATCATCATCATTGAGTACTGTATCAGCTATATCTA
TTTCTTCAGGAACACTGTGTGATTTCATAGATGAAAGCAGTTCCATTACAGGAATCACTTCTCCAGTAAGCATTGGA
ATAATACTCTGACCCTGATCTTCCATTCTCTCTGTGCCTTCTAAGATAATCTTCTGGACATTTTCTTGTCTTTCCTT
ATGCATCCATATTCTTCCTTTCCGGATTATATGTGTTAATCTATCAACACACACTCTATGAAGTGGGAGGTAGAAAC

TAAGTTCTGTCTGTGGAAGTATAATTGATAGTCCATATGTGCTGCGATCCCCATTCCAGTTTCCATCAAAGATTAAT
GAAACAATAATCACTCCCTTTTCAGACAAGACAAAAAACTTTACATCTATAGCACCACTCTCTGCATTTCGAAGGAT
TTCTCCATTTAGAGTGTGGTTGGCAAGAAAAGTTATTTCTCCATCACTGAGAAGTACCTGTTCTGTCTTTGGAGCCC
AAATGTGCCTTACTCTAGGACCAAGAATATTGTCCCAGTAAGCAAAAGTAGCTGCTAATAAAGGTGATTTGCCACTT
AAAGCAATCTCTGTCTTGGCAACAGCTGGAGATGGCGGTGGGCAAAGAGTCGACATCACTGCATTCCAACTGTCACA
TTATCCAAATGCTCCGGAGATATCCACTCGCCACCGCCTGCGCCTCCGCCGCCGCGGGCGCAGGCACCGCAACC

SEQ ID NO:133
>NC_000009.12:c27573866-27546546 智人 (Homo sapiens) 染色体 9, GRCh38.p13 初级组件;携带 C9orf72 基因的人染色体 9 的一部分 (染色体 9 组件的核苷酸 27546546..27573866)
ACGTAACCTACGGTGTCCCGCTAGGAAAGAGAGGTGCGTCAAACAGCGACAAGTTCCGCCCACGTAAAAG
ATGACGCTTGGTGTGTCAGCCGTCCCTGCTGCCCGGTTGCTTCTCTTTTGGGGGCGGGGTCTAGCAAGAG
CAGGTGTGGGTTTAGGAGGTGTGTGTTTTTGTTTTTCCCACCCTCTCTCCCCACTACTTGCTCTCACAGT
ACTCGCTGAGGGTGAACAAGAAAAGACCTGATAAAGATTAACCAGAAGAAAACAAGGAGGGAAACAACCG
CAGCCTGTAGCAAGCTCTGGAACTCAGGAGTCGCGCGCTAGGGGCCGGGGCCGGGGCCGGGGCGTGGTCG
GGGCGGGCCCGGGGGCGGGCCCGGGGCGGGGCTGCGGTTGCGGTGCCTGCGCCCGCGGCGGCGGAGGCGC
AGGCGGTGGCGAGTGGGTGAGTGAGGAGGCGGCATCCTGGCGGGTGGCTGTTTGGGGTTCGGCTGCCGGG
AAGAGGCGCGGGTAGAAGCGGGGGCTCTCCTCAGAGCTCGACGCATTTTTACTTTCCCTCTCATTTCTCT
GACCGAAGCTGGGTGTCGGGCTTTCGCCTCTAGCGACTGGTGGAATTGCCTGCATCCGGGCCCCGGGCTT
CCCGGCGGCGGCGGCGGCGGCGGCGGCGCAGGGACAAGGGATGGGGATCTGGCCTCTTCCTTGCTTTCCC
GCCCTCAGTACCCGAGCTGTCTCCTTCCCGGGGACCCGCTGGGAGCGCTGCCGCTGCGGGCTCGAGAAAA
GGGAGCCTCGGGTACTGAGAGGCCTCGCCTGGGGGAAGGCCGGAGGGTGGGCGGCGCGCGGCTTCTGCGG
ACCAAGTCGGGGTTCGCTAGGAACCCGAGACGGTCCCTGCCGGCGAGGAGATCATGCGGGATGAGATGGG
GGTGTGGAGACGCCTGCACAATTTCAGCCCAAGCTTCTAGAGAGTGGTGATGACTTGCATATGAGGGCAG
CAATGCAAGTCGGTGTGCTCCCCATTCTGTGGGACATGACCTGGTTGCTTCACAGCTCCGAGATGACACA
GACTTGCTTAAAGGAAGTGACTATTGTGACTTGGGCATCACTTGACTGATGGTAATCAGTTGTCTAAAGA
AGTGCACAGATTACATGTCCGTGTGCTCATTGGGTCTATCTGGCCGCGTTGAACACCACCAGGCTTTGTA
TTCAGAAACAGGAGGGAGGTCCTGCACTTTCCCAGGAGGGGTGGCCCTTTCAGATGCAATCGAGATTGTT
AGGCTCTGGGAGAGTAGTTGCCTGGTTGTGGCAGTTGGTAAATTTCTATTCAAACAGTTGCCATGCACCA
GTTGTTCACAACAAGGGTACGTAATCTGTCTGGCATTACTTCTACTTTTGTACAAAGGATCAAAAAAAAA
AAAGATACTGTTAAGATATGATTTTTCTCAGACTTTGGGAAACTTTTAACATAATCTGTGAATATCACAG
[1130] AAACAAGACTATCATATAGGGGATATTAATAACCTGGAGTCAGAATACTTGAAATACGGTGTCATTTGAC
ACGGGCATTGTTGTCACCACCTCTGCCAAGGCCTGCCACTTTAGGAAAACCCTGAATCAGTTGGAAACTG
CTACATGCTGATAGTACATCTGAAACAAGAACGAGAGTAATTACCACATTCCAGATTGTTCACTAAGCCA
GCATTTACCTGCTCCAGGAAAAAATTACAAGCACCTTATGAAGTTGATAAAATATTTTGTTTGGCTATGT
TGGCACTCCACAATTTGCTTTCAGAGAAACAAAGTAAACCAAGGAGGACTTCTGTTTTTCAAGTCTGCCC
TCGGGTTCTATTCTACGTTAATTAGATAGTTCCCAGGAGGACTAGGTTAGCCTACCTATTGTCTGAGAAA
CTTGGAACTGTGAGAAATGGCCAGATAGTGATATGAACTTCACCTTCCAGTCTTCCCTGATGTTGAAGAT
TGAGAAAGTGTTGTGAACTTTCTGGTACTGTAAACAGTTCACTGTCCTTGAAGTGGTCCTGGGCAGCTCC
TGTTGTGGAAAGTGGACGGTTTAGGATCCTGCTTCTCTTTGGGCTGGGAGAAAATAAACAGCATGGTTAC
AAGTATTGAGAGCCAGGTTGGAGAAGGTGGCTTACACCTGTAATGCCAGAGCTTTGGGAGGCGGAGGCAA
GAGGATCACTTGAAGCCAGGAGTTCAAGCTCAACCTGGGCAACGTAGACCCTGTCTCTACAAAAAATTAA
AAACTTAGCCGGGCGTGGTGATGTGCACCTGTAGTCCTAGCTACTTGGGAGGCTGAGGCAGGAGGGTCAT
TTGAGCCCAAGAGTTTGAAGTTACCGAGAGCTATGATCCTGCCAGTGCATTCCAGCCTGGATGACAAAAC
GAGACCCTGTCTCTAAAAAACAAGAAGTGAGGGCTTTATGATTGTAGAATTTTCACTACAATAGCAGTGG
ACCAACCACCTTTCTAAATACCAATCAGGGAAGAGATGGTTGATTTTTTAACAGACGTTTAAAGAAAAAG
CAAAACCTCAAACTTAGCACTCTACTAACAGTTTTAGCAGATGTTAATTAATGTAATCATGTCTGCATGT
ATGGGATTATTTCCAGAAAGTGTATTGGGAAACCTCTCATGAACCCTGTGAGCAAGCCACCGTCTCACTC
AATTTGAATCTTGGCTTCCCTCAAAAGACTGGCTAATGTTTGGTAACTCTCTGGAGTAGACAGCACTACA
TGTACGTAAGATAGGTACATAAACAACTATTGGTTTTGAGCTGATTTTTTTCAGCTGCATTTGCATGTAT
GGATTTTTCTCACCAAAGACGATGACTTCAAGTATTAGTAAAATAATTGTACAGCTCTCCTGATTATACT
TCTCTGTGACATTTCATTTCCCAGGCTATTTCTTTTGGTAGGATTTAAAACTAAGCAATTCAGTATGATC
TTTGTCCTTCATTTTCTTTCTTATTCTTTTTGTTTGTTTGTTTGTTTGTTTTTTTCTTGAGGCAGAGTCT
CTCTCTGTCGCCCAGGCTGGAGTGCAGTGGCGCCATCTCAGCTCATTGCAACCTCTGCCACCTCCGGGTT
CAAGAGATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGTGTCCACCACCACACCCGGCTAAT
TTTTTGTATTTTTAGTAGAGGTGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAGCTCCTGACCTCAGG
TGATCCACCTGCCTCGGCCTACCAAAGAGCTGGGATAACAGGTGTGACCCACCATGCCCGGCCCATTTTT
TTTTTCTTATTCTGTTAGGAGTGAGAGTGTAACTAGCAGTATAATAGTTCAATTTTCACAACGTGGTAAA
AGTTTCCCTATAATTCAATCAGATTTTGCTCCAGGGTTCAGTTCTGTTTTAGGAAATACTTTTATTTTCA
GTTTAATGATGAAATATTAGAGTTGTAATATTGCCTTTATGATTATCCACCTTTTTAACCTAAAAGAATG
AAAGAAAAATATGTTTGCAATATAATTTTATGGTTGTATGTTAACTTAATTCATTATGTTGGCCTCCAGT
TTGCTGTTGTTAGTTATGACAGCAGTAGTGTCATTACCATTTCAATTCAGATTACATTCCTATATTTGAT
CATTGTAAACTGACTGCTTACATTGTATTAAAAACAGTGGATATTTTAAAGAAGCTGTACGGCTTATATC
TAGTGCTGTCTCTTAAGACTATTAAATTGATACAACATATTTAAAAGTAAATATTACCTAAATGAATTTT

TGAAATTACAAATACACGTGTTAAAACTGTCGTTGTGTTCAACCATTTCTGTACATACTTAGAGTTAACT
GTTTTGCCAGGCTCTGTATGCCTACTCATAATATGATAAAAGCACTCATCTAATGCTCTGTAAATAGAAG
TCAGTGCTTTCCATCAGACTGAACTCTCTTGACAAGATGTGGATGAAATTCTTTAAGTAAAATTGTTTAC
TTTGTCATACATTTACAGATCAAATGTTAGCTCCCAAAGCAATCATATGGCAAAGATAGGTATATCATAG
TTTGCCTATTAGCTGCTTTGTATTGCTATTATTATAAATAGACTTCACAGTTTTAGACTTGCTTAGGTGA
AATTGCAATTCTTTTTACTTTCAGTCTTAGATAACAAGTCTTCAATTATAGTACAATCACACATTGCTTA
GGAATGCATCATTAGGCGATTTTGTCATTATGCAAACATCATAGAGTGTACTTACACAAACCTAGATAGT
ATAGCCTTTATGTACCTAGGCCGTATGGTATAGTCTGTTGCTCCTAGGCCACAAACCTGTACAACTGTTA
CTGTACTGAATACTATAGACAGTTGTAACACAGTGGTAAATATTTATCTAAATATATGCAAACAGAGAAA
AGGTACAGTAAAAGTATGGTATAAAAGATAATGGTATACCTGTGTAGGCCACTTACCACGAATGGAGCTT
GCAGGACTAGAAGTTGCTCTGGGTGAGTCAGTGAGTGAGTGGTGAATTAATGTGAAGGCCTAGAACACTG
TACACCACTGTAGACTATAAACACAGTACGCTGAAGCTACACCAAATTTATCTTAACAGTTTTTCTTCAA
TAAAAAATTATAACTTTTTAACTTTGTAAACTTTTTAATTTTTTAACTTTTAAAATACTTAGCTTGAAAC
ACAAATACATTGTATAGCTATACAAAAATATTTTTTCTTTGTATCCTTATTCTAGAAGCTTTTTTCTATT
TTCTATTTTAAATTTTTTTTTTTACTTGTTAGTCGTTTTTGTTAAAAACTAAAACACACACACTTTCACC
TAGGCATAGACAGGATTAGGATCATCAGTATCACTCCCTTCCACCTCACTGCCTTCCACCTCCACATCTT
GTCCCACTGGAAGGTTTTTAGGGGCAATAACACACATGTAGCTGTCACCTATGATAACAGTGCTTTCTGT
TGAATACCTCCTGAAGGACTTGCCTGAGGCTGTTTTACATTTAACTTAAAAAAAAAAAAAGTAGAAGGAG
TGCACTCTAAAATAACAATAAAAGGCATAGTATAGTGAATACATAAACCAGCAATGTAGTAGTTTATTAT
CAAGTGTTGTACACTGTAATAATTGTATGTGCTATACTTTAAATAACTTGCAAAATAGTACTAAGACCTT
ATGATGGTTACAGTGTCACTAAGGCAATAGCATATTTTCAGGTCCATTGTAATCTAATGGGACTACCATC
ATATATGCAGTCTACCATTGACTGAAACGTTACATGGCACATAACTGTATTTGCAAGAATGATTTGTTTT
ACATTAATATCACATAGGATGTACCTTTTTAGAGTGGTATGTTTATGTGGATTAAGATGTACAAGTTGAG
CAAGGGGACCAAGAGCCCTGGGTTCTGTCTTGGATGTGAGCGTTTATGTTCTTCTCCTCATGTCTGTTTT
CTCATTAAATTCAAAGGCTTGAACGGGCCCTATTTAGCCCTTCTGTTTTCTACGTGTTCTAAATAACTAA
AGCTTTTAAATTCTAGCCATTTAGTGTAGAACTCTCTTTGCAGTGATGAAATGCTGTATTGGTTTCTTGG
CTAGCATATTAAATATTTTTATCTTTGTCTTGATACTTCAATGTCGTTTTAAACATCAGGATCGGGCTTC
AGTATTCTCATAACCAGAGAGTTCACTGAGGATACAGGACTGTTTGCCCATTTTTTGTTATGGCTCCAGA
CTTGTGGTATTTCCATGTCTTTTTTTTTTTTTTTTTTTTGACCTTTTAGCGGCTTTAAAGTATTTCTGT
TGTTAGGTGTTGTATTACTTTTCTAAGATTACTTAACAAAGCACCACAAACTGAGTGGCTTTAAACAACA
GCAATTTATTCTCTCACAATTCTAGAAGCTAGAAGTCCGAAATCAAAGTGTTGACAGGGGCATGATCTTC
AAGAGAGAAGACTCTTTCCTTGCCTCTTCCTGGCTTCTGGTGGTTACCAGCAATCCTGAGTGTTCCTTTC
[1131] TTGCCTTGTAGTTTCAACAATCCAGTATCTGCCTTTTGTCTTCACATGGCTGTCTACCATTTGTCTCTGT
GTCTCCAAATCTCTCTCCTTATAAACACAGCAGTTATTGGATTAGGCCCCACTCTAATCCAGTATGACCC
CATTTTAACATGATTACACTTATTTCTAGATAAGGTCACATTCACGTACACCAAGGGTTAGGAATTGAAC
ATATCTTTTTGGGGGACACAATTCAACCCACAAGTGTCAGTCTCTAGCTGAGCCTTTCCCTTCCTGTTTT
TCTCCTTTTTAGTTGCTATGGGTTAGGGGCCAAATCTCCAGTCATACTAGAATTGCACATGGACTGGATA
TTTGGGAATACTGCGGGTCTATTCTATGAGCTTTAGTATGTAACATTTAATATCAGTGTAAAGAAGCCCT
TTTTTAAGTTATTTCTTTGAATTTCTAAATGTATGCCCTGAATATAAGTAACAAGTTACCATGTCTTGTA
AAATGATCATATCAACAAACATTTAATGTGCACCTACTGTGCTAGTTGAATGTCTTTATCCTGATAGGAG
ATAACAGGATTCCACATCTTTGACTTAAGAGGACAAACCAAATATGTCTAAATCATTTGGGGTTTTGATG
GATATCTTTAAATTGCTGAACCTAATCATTGGTTTCATATGTCATTGTTTAGATATCTCCGGAGCATTTG
GATAATGTGACAGTTGGAATGCAGTGATGTCGACTCTTTGCCCACCGCCATCTCCAGCTGTTGCCAAGAC
AGAGATTGCTTTAAGTGGCAAATCACCTTTATTAGCAGCTACTTTTGCTTACTGGGACAATATTCTTGGT
CCTAGAGTAAGGCACATTTGGGCTCCAAAGACAGAACAGGTACTTCTCAGTGATGGAGAAATAACTTTTC
TTGCCAACCACACTCTAAATGGAGAAATCCTTCGAAATGCAGAGAGTGGTGCTATAGATGTAAAGTTTTT
TGTCTTGTCTGAAAAGGGAGTGATTATTGTTTCATTAATCTTTGATGGAAACTGGAATGGGGATCGCAGC
ACATATGGACTATCAATTATACTTCCACAGACAGAACTTAGTTTCTACCTCCCACTTCATAGAGTGTGTG
TTGATAGATTAACACATATAATCCGGAAAGGAAGAATATGGATGCATAAGGTAAGTGATTTTTCAGCTTA
TTAATCATGTTAACCTATCTGTTGAAAGCTTATTTTCTGGTACATATAAATCTTATTTTTTTAATTATAT
GCAGTGAACATCAAACAATAAATGTTATTTATTTTGCATTTACCCTATTAGATACAAATACATCTGGTCT
GATACCTGTCATCTTCATATTAACTGTGGAAGGTACGAAATGGTAGCTCCACATTATAGATGAAAAGCTA
AAGCTTAGACAAATAAAGAAACTTTTAGACCCTGGATTCTTCTTGGGAGCCTTTGACTCTAATACCTTTT
GTTTCCCTTTCATTGCACAATTCTGTCTTTTGCTTACTACTATGTGTAAGTATAACAGTTCAAAGTAATA
GTTTCATAAGCTGTTGGTCATGTAGCCTTTGGTCTCTTTAACCTCTTTGCCAAGTTCCCAGGTTCATAAA
ATGAGGAGGTTGAATGGAATGGTTCCCAAGAGAATTCCTTTTAATCTTACAGAAATTATTGTTTTCCTAA
ATCCTGTAGTTGAATATATAATGCTATTTACATTTCAGTATAGTTTTGATGTATCTAAAGAACACATTGA
ATTCTCCTTCCTGTGTTCCAGTTTGATACTAACCTGAAAGTCCATTAAGCATTACCAGTTTTAAAAGGCT
TTTGCCCAATAGTAAGGAAAAATAATATCTTTTAAAAGAATAATTTTTTACTATGTTTGCAGGCTTACTT
CCTTTTTTCTCACATTATGAAACTCTTAAAATCAGGAGAATCTTTTAAACAACATCATAATGTTTAATTT
GAAAAGTGCAAGTCATTCTTTTCCTTTTTGAAACTATGCAGATGTTACATTGACTGTTTTCTGTGAAGTT
ATCTTTTTTTCACTGCAGAATAAAGGTTGTTTTGATTTTATTTTGTATTGTTTATGAGAACATGCATTTG
TTGGGTTAATTTCCTACCCCTGCCCCCATTTTTTCCCTAAAGTAGAAAGTATTTTTCTTGTGAACTAAAT
TACTACACAAGAACATGTCTATTGAAAAATAAGCAAGTATCAAAATGTTGTGGGTTGTTTTTTTAAATAA
ATTTTCTCTTGCTCAGGAAAGACAAGAAAATGTCCAGAAGATTATCTTAGAAGGCACAGAGAGAATGGAA
GATCAGGTATATGCAAATTGCATACTGTCAAATGTTTTTCTCACAGCATGTATCTGTATAAGGTTGATGG

CTACATTTGTCAAGGCCTTGGAGACATACGAATAAGCCTTTAATGGAGCTTTTATGGAGGTGTACAGAAT
AAACTGGAGGAAGATTTCCATATCTTAAACCCAAAGAGTTAAATCAGTAAACAAAGGAAAATAGTAATTG
CATCTACAAATTAATATTTGCTCCCTTTTTTTTTCTGTTTGCCCAGAATAAATTTTGGATAACTTGTTCA
TAGTAAAAATAAAAAAAATTGTCTCTGATATGTTCTTTAAGGTACTACTTCTCGAACCTTTCCCTAGAAG
TAGCTGTAACAGAAGGAGAGCATATGTACCCCTGAGGTATCTGTCTGGGGTGTAGGCCCAGGTCCACACA
ATATTTCTTCTAAGTCTTATGTTGTATCGTTAAGACTCATGCAATTTACATTTTATTCCATAACTATTTT
AGTATTAAAATTTGTCAGTGATATTTCTTACCCTCTCCTCTAGGAAAATGTGCCATGTTTATCCCTTGGC
TTTGAATGCCCCTCAGGAACAGACACTAAGAGTTTGAGAAGCATGGTTACAAGGGTGTGGCTTCCCCTGC
GGAAACTAAGTACAGACTATTTCACTGTAAAGCAGAGAAGTTCTTTTGAAGGAGAATCTCCAGTGAAGAA
AGAGTTCTTCACTTTTACTTCCATTTCCTCTTGTGGGTGACCCTCAATGCTCCTTGTAAAACTCCAATAT
TTTAAACATGGCTGTTTTGCCTTTCTTTGCTTCTTTTTAGCATGAATGAGACAGATGATACTTTAAAAAA
GTAATTAAAAAAAAAAACTTGTGAAAATACATGGCCATAATACAGAACCCAATACAATGATCTCCTTTAC
CAAATTGTTATGTTTGTACTTTTGTAGATAGCTTTCCAATTCAGAGACAGTTATTCTGTGTAAAGGTCTG
ACTTAACAAGAAAAGATTTCCCTTTACCCAAAGAATCCCAGTCCTTATTTGCTGGTCAATAAGCAGGGTC
CCCAGGAATGGGGTAACTTTCAGCACCCTCTAACCCACTAGTTATTAGTAGACTAATTAAGTAAACTTAT
CGCAAGTTGAGGAAACTTAGAACCAACTAAAATTCTGCTTTTACTGGGATTTTGTTTTTTCAAACCAGAA
ACCTTTACTTAAGTTGACTACTATTAATGAATTTTGGTCTCTCTTTTAAGTGCTCTTCTTAAAAATGTTA
TCTTACTGCTGAGAAGTTCAAGTTTGGGAAGTACAAGGAGGAATAGAAACTTAAGAGATTTTCTTTTAGA
GCCTCTTCTGTATTTAGCCCTGTAGGATTTTTTTTTTTTTTTTTTTTTTGGTGTTGTTGAGCTTCAGTG
AGGCTATTCATTCACTTATACTGATAATGTCTGAGATACTGTGAATGAAATACTATGTATGCTTAAACCT
AAGAGGAAATATTTTCCCAAAATTATTCTTCCCGAAAAGGAGGAGTTGCCTTTTGATTGAGTTCTTGCAA
ATCTCACAACGACTTTATTTTGAACAATACTGTTTGGGGATGATGCATTAGTTTGAAACAACTTCAGTTG
TAGCTGTCATCTGATAAAATTGCTTCACAGGGAAGGAAATTTAACACGGATCTAGTCATTATTCTTGTTA
GATTGAATGTGTGAATTGTAATTGTAAACAGGCATGATAATTATTACTTTAAAAACTAAAAACAGTGAAT
AGTTAGTTGTGGAGGTTACTAAAGGATGGTTTTTTTTTAAATAAAACTTTCAGCATTATGCAAATGGGCA
TATGGCTTAGGATAAAACTTCCAGAAGTAGCATCACATTTAAATTCTCAAGCAACTTAATAATATGGGGC
TCTGAAAAACTGGTTAAGGTTACTCCAAAAATGGCCCTGGGTCTGACAAAGATTCTAACTTAAAGATGCT
TATGAAGACTTTGAGTAAAATCATTTCATAAAATAAGTGAGGAAAAACAACTAGTATTAAATTCATCTTA
AATAATGTATGATTTAAAAAATATGTTTAGCTAAAAATGCATAGTCATTTGACAATTTCATTTATATCTC
AAAAAATTTACTTAACCAAGTTGGTCACAAAACTGATGAGACTGGTGGTGGTAGTGAATAAATGAGGGAC
CATCCATATTTGAGACACTTTACATTTGTGATGTGTTATACTGAATTTTCAGTTTGATTCTATAGACTAC
AAATTTCAAAATTACAATTTCAAGATGTAATAAGTAGTAATATCTTGAAATAGCTCTAAAGGGAATTTTT
[1132] CTGTTTTATTGATTCTTAAAATATATGTGCTGATTTTGATTTGCATTTGGGTAGATTATACTTTTATGAG
TATGGAGGTTAGGTATTGATTCAAGTTTTCCTTACCTATTTGGTAAGGATTTCAAAGTCTTTTTGTGCTT
GGTTTTCCTCATTTTTAAATATGAAATATATTGATGACCTTTAACAAATTTTTTTTATCTCAAATTTTAA
AGGAGATCTTTTCTAAAAGAGGCATGATGACTTAATCATTGCATGTAACAGTAAACGATAAACCAATGAT
TCCATACTCTCTAAAGAATAAAAGTGAGCTTTAGGGCCGGGCATGGTCAGAAATTTGACACCAACCTGGC
CAACATGGCGAAACCCCGTCTCTACTAAAAATACAAAAATCAGCCGGGCATGGTGGCGGCACCTATAGTC
CCAGCTACTTGGGAGGATGAGACAGGAGAGTCACTTGAACCTGGGAGGAGAGGTTGCAGTGAGCTGAGAT
CACGCCATTGCACTCCAGCCTGAGCAATGAAAGCAAAACTCCATCTCAAAAAAAAAAAAAAGAAAAGAAAG
AATAAAAGTGAGCTTTGGATTGCATATAAATCCTTTAGACATGTAGTAGACTTGTTTGATACTGTGTTTG
AACAAATTACGAAGTATTTTCATCAAAGAATGTTATTGTTTGATGTTATTTTTATTTTTTATTGCCCAGC
TTCTCTCATATTACGTGATTTTCTTCACTTCATGTCACTTTATTGTGCAGGGTCAGAGTATTATTCCAAT
GCTTACTGGAGAAGTGATTCCTGTAATGGAACTGCTTTCATCTATGAAATCACACAGTGTTCCTGAAGAA
ATAGATGTAAGTTTAAATGAGAGCAATTATACACTTTATGAGTTTTTTGGGGTTATAGTATTATTATGTA
TATTATTAATATTCTAATTTTAATAGTAAGGACTTTGTCATACATACTATTCACATACAGTATTAGCCAC
TTTAGCAAATAAGCACACACAAAATCCTGGATTTTATGGCAAAACAGAGGCATTTTTGATCAGTGATGAC
AAAATTAAATTCATTTTGTTTATTTCATTACTTTTATAATTCCTAAAAGTGGGAGGATCCCAGCTCTTAT
AGGAGCAATTAATATTTAATGTAGTGTCTTTTGAAACAAAACTGTGTGCCAAAGTAGTAACCATTAATGG
AAGTTTACTTGTAGTCACAAATTTAGTTTCCTTAATCATTTGTTGAGGACGTTTTGAATCACACACTATG
AGTGTTAAGAGATACCTTTAGGAAACTATTCTTGTTGTTTTCTGATTTTGTCATTTAGGTTAGTCTCCTG
ATTCTGACAGCTCAGAAGAGGAAGTTGTTCTTGTAAAAATTGTTTAACCTGCTTGACCAGCTTTCACATT
TGTTCTTCTGAAGTTTATGGTAGTGCACAGAGATTGTTTTTTGGGGAGTCTTGATTCTCGGAAATGAAGG
CAGTGTGTTATATTGAATCCAGACTTCCGAAAACTTGTATATTAAAAGTGTTATTTCAACACTATGTTAC
AGCCAGACTAATTTTTTTATTTTTTGATGCATTTTAGATAGCTGATACAGTACTCAATGATGATGATATT
GGTGACAGCTGTCATGAAGGCTTTCTTCTCAAGTAAGAATTTTTCTTTTCATAAAAGCTGGATGAAGCAG
ATACCATCTTATGCTCACCTATGACAAGATTTGGAAGAAAGAAAATAACAGACTGTCTACTTAGATTGTT
CTAGGGACATTACGTATTTGAACTGTTGCTTAAATTTGTGTTATTTTTCACTCATTATATTTCTATATAT
ATTTGGTGTTATTCCATTTGCTATTTAAAGAAACCGAGTTTCCATCCCAGACAAGAAATCATGGCCCCTT
GCTTGATTCTGGTTTCTTGTTTTACTTCTCATTAAAGCTAACAGAATCCTTTCATATTAAGTTGTACTGT
AGATGAACTTAAGTTATTTAGGCGTAGAACAAAATTATTCATATTTATACTGATCTTTTTCCATCCAGCA
GTGGAGTTTAGTACTTAAGAGTTTGTGCCCTTAAACCAGACTCCCTGGATTAATGCTGTGTACCCGTGGG
CAAGGTGCCTGAATTCTCTATACACCTATTTCCTCATCTGTAAAATGGCAATAATAGTAATAGTACCTAA
TGTGTAGGGTTGTTATAAGCATTGAGTAAGATAAATAATATAAAGCACTTAGAACAGTGCCTGGAACATA
AAAACACTTAATAATAGCTCATAGCTAACATTTCCTATTTACATTTCTTCTAGAAATAGCCAGTATTTGT
TGAGTGCCTACATGTTAGTTCCTTTACTAGTTGCTTTACATGTATTATCTTATATTCTGTTTTAAAGTTT

CTTCACAGTTACAGATTTTCATGAAATTTTACTTTTAATAAAAGAGAAGTAAAAGTATAAAGTATTCACT
TTTATGTTCACAGTCTTTTCCTTTAGGCTCATGATGGAGTATCAGAGGCATGAGTGTGTTTAACCTAAGA
GCCTTAATGGCTTGAATCAGAAGCACTTTAGTCCTGTATCTGTTCAGTGTCAGCCTTTCATACATCATTT
TAAATCCCATTTGACTTTAAGTAAGTCACTTAATCTCTCTACATGTCAATTTCTTCAGCTATAAAATGAT
GGTATTTCAATAAATAAATACATTAATTAAATGATATTATACTGACTAATTGGGCTGTTTTAAGGCTCAA
TAAGAAAATTTCTGTGAAAGGTCTCTAGAAAATGTAGGTTCCTATACAAATAAAAGATAACATTGTGCTT
ATAGCTTCGGTGTTTATCATATAAAGCTATTCTGAGTTATTTGAAGAGCTCACCTACTTTTTTTTGTTTT
TAGTTTGTTAAATTGTTTTATAGGCAATGTTTTTAATCTGTTTTCTTTAACTTACAGTGCCATCAGCTCA
CACTTGCAAACCTGTGGCTGTTCCGTTGTAGTAGGTAGCAGTGCAGAGAAAGTAAATAAGGTAGTTTATT
TTATAATCTAGCAAATGATTTGACTCTTTAAGACTGATGATATATCATGGATTGTCATTTAAATGGTAGG
TTGCAATTAAAATGATCTAGTAGTATAAGGAGGCAATGTAATCTCATCAAATTGCTAAGACACCTTGTGG
CAACAGTGAGTTTGAAATAAACTGAGTAAGAATCATTTATCAGTTTATTTTGATAGCTCGGAAATACCAG
TGTCAGTAGTGTATAAATGGTTTTGAGAATATATTAAAATCAGATATATAAAAAAAATTACTCTTCTATT
TCCCAATGTTATCTTTAACAAATCTGAAGATAGTCATGTACTTTTGGTAGTAGTTCCAAAGAAATGTTAT
TTGTTTATTCATCTTGATTTCATTGTCTTCGCTTTCCTTCTAAATCTGTCCCTTCTAGGGAGCTATTGGG
ATTAAGTGGTCATTGATTATTATACTTTATTCAGTAATGTTTCTGACCCTTTCCTTCAGTGCTACTTGAG
TTAATTAAGGATTAATGAACAGTTACATTTCCAAGCATTAGCTAATAAACTAAAGGATTTTGCACTTTTC
TTCACTGACCATTAGTTAGAAAGAGTTCAGAGATAAGTATGTGTATCTTTCAATTTCAGCAAACCTAATT
TTTTAAAAAAAGTTTTACATAGGAAATATGTTGGAAATGATACTTTACAAAGATATTCATAATTTTTTTT
TGTAATCAGCTACTTTGTATATTTACATGAGCCTTAATTTATATTTCTCATATAACCATTTATGAGAGCT
TAGTATACCTGTGTCATTATATTGCATCTACGAACTAGTGACCTTATTCCTTCTGTTACCTCAAACAGGT
GGCTTTCCATCTGTGATCTCCAAAGCCTTAGGTTGCACAGAGTGACTGCCGAGCTGCTTTATGAAGGGAG
AAAGGCTCCATAGTTGGAGTGTTTTTTTTTTTTTTTAAACATTTTTCCCATCCTCCATCCTCTTGAGG
GAGAATAGCTTACCTTTTATCTTGTTTTAATTTGAGAAAGAAGTTGCCACCACTCTAGGTTGAAAACCAC
TCCTTTAACATAATAACTGTGGATATGGTTTGAATTTCAAGATAGTTACATGCCTTTTTATTTTTCCTAA
TAGAGCTGTAGGTCAAATATTATTAGAATCAGATTTCTAAATCCCACCCAATGACCTGCTTATTTTAAAT
CAAATTCAATAATTAATTCTCTTCTTTTTGGAGGATCTGGACATTCTTTGATATTTCTTACAACGAATTT
CATGTGTAGACCCACTAAACAGAAGCTATAAAAGTTGCATGGTCAAATAAGTCTGAGAAAGTCTGCAGAT
GATATAATTCACCTGAAGAGTCACAGTATGTAGCCAAATGTTAAAGGTTTTGAGATGCCATACAGTAAAT
TTACCAAGCATTTTCTAAATTTATTTGACCACAGAATCCCTATTTTAAGCAACAACTGTTACATCCCATG
GATTCCAGGTGACTAAAGAATACTTATTTCTTAGGATATGTTTTATTGATAATAACAATTAAAATTTCAG
ATATCTTTCATAAGCAAATCAGTGGTCTTTTTACTTCATGTTTTAATGCTAAAATATTTTCTTTTATAGA
[1133] TAGTCAGAACATTATGCCTTTTTCTGACTCCAGCAGAGAGAAAATGCTCCAGGTTATGTGAAGCAGAATC
ATCATTTAAATATGAGTCAGGGCTCTTTGTACAAGGCCTGCTAAAGGTATAGTTTCTAGTTATCACAAGT
GAAACCACTTTTCTAAAATCATTTTTGAGACTCTTTATAGACAAATCTTAAATATTAGCATTTAATGTAT
CTCATATTGACATGCCCAGAGACTGACTTCCTTTACACAGTTCTGCACATAGACTATATGTCTTATGGAT
TTATAGTTAGTATCATCAGTGAAACACCATAGAATACCCTTTGTGTTCCAGGTGGGTCCCTGTTCCTACA
TGTCTAGCCTCAGGACTTTTTTTTTTTTTAACACATGCTTAAATCAGGTTGCACATCAAAAATAAGATCAT
TTCTTTTTAACTAAATAGATTTGAATTTTATTGAAAAAAAATTTTAAACATCTTTAAGAAGCTTATAGGA
TTTAAGCAATTCCTATGTATGTGTACTAAAATATATATATTTCTATATATAATATATTAGAAAAAAAT
TGTATTTTTCTTTTATTTGAGTCTACTGTCAAGGAGCAAAACAGAGAAATGTAAATTAGCAATTATTTAT
AATACTTAAAGGGAAGAAAGTTGTTCACCTTGTTGAATCTATTATTGTTATTTCAATTATAGTCCCAAGA
CGTGAAGAAATAGCTTTCCTAATGGTTATGTGATTGTCTCATAGTGACTACTTTCTTGAGGATGTAGCCA
CGGCAAAATGAAATAAAAAAATTTAAAAATTGTTGCAAATACAAGTTATATTAGGCTTTTGTGCATTTTC
AATAATGTGCTGCTATGAACTCAGAATGATAGTATTTAAATATAGAAACTAGTTAAAGGAAACGTAGTTT
CTATTTGAGTTATACATATCTGTAAATTAGAACTTCTCCTGTTAAAGGCATAATAAAGTGCTTAATACTT
TTGTTTCCTCAGCACCCTCTCATTTAATTATATAATTTTAGTTCTGAAAGGGACCTATACCAGATGCCTA
GAGGAAATTTCAAAACTATGATCTAATGAAAAAATATTTAATAGTTCTCCATGCAAATACAAATCATATA
GTTTTCCAGAAAATACCTTTGACATTATACAAAGATGATTATCACAGCATTATAATAGTAAAAAAATGGA
AATAGCCTCTTTCTTCTGTTCTGTTCATAGCACAGTGCCTCATACGCAGTAGGTTATTATTACATGGTAA
CTGGCTACCCCAACTGATTAGGAAAGAAGTAAATTTGTTTTATAAAAATACATACTCATTGAGGTGCATA
GAATAATTAAGAAATTAAAAGACACTTGTAATTTTGAATCCAGTGAATACCCACTGTTAATATTTGGTAT
ATCTCTTTCTAGTCTTTTTTTCCCTTTTGCATGTATTTTCTTTAAGACTCCCACCCCCACTGGATCATCT
CTGCATGTTCTAATCTGCTTTTTTCACAGCAGATTCTAAGCCTCTTTGAATATCAACACAAACTTCAACA
ACTTCATCTATAGATGCCAAATAATAAATTCATTTTTATTTACTTAACCACTTCCTTTGGATGCTTAGGT
CATTCTGATGTTTTGCTATTGAAACCAATGCTATACTGAACACTTCTGTCACTAAAACTTTGCACACACT
CATGAATAGCTTCTTAGGATAAATTTTTAGAGATGGATTTGCTAAATCAGAGACCATTTTTTAAAATTAA
AAAACAATTATTCATATCGTTTGGCATGTAAGACAGTAAATTTTCCTTTTATTTTGACAGGATTCAACTG
GAAGCTTTGTGCTGCCTTTCCGGCAAGTCATGTATGCTCCATATCCCACCACACACATAGATGTGGATGT
CAATACTGTGAAGCAGATGCCACCCTGTCATGAACATATTTATAATCAGCGTAGATACATGAGATCCGAG
CTGACAGCCTTCTGGAGAGCCACTTCAGAAGAAGACATGGCTCAGGATACGATCATCTACACTGACGAAA
GCTTTACTCCTGATTTGTACGTAATGCTCTGCCTGCTGGTACTGTAGTCAAGCAATATGAAATTGTGTCT
TTTACGAATAAAAACAAAACAGAAGTTGCATTTAAAAAGAAAGAAATATTACCAGCAGAATTATGCTTGA
AGAAACATTTAATCAAGCATTTTTTTCTTAAATGTTCTTCTTTTTCCATACAATTGTGTTTACCCTAAAA
TAGGTAAGATTAACCCTTAAAGTAAATATTTAACTATTTGTTTAATAAATATATATTGAGCTCCTAGGCA
CTGTTCTAGGTACCGGGCTTAATAGTGGCCAACCAGACAGCCCCAGCCCCAGCCCCTACATTGTGTATAG

TCTATTATGTAACAGTTATTGAATGGACTTATTAACAAAACCAAAGAAGTAATTCTAAGTCTTTTTTTTC
TTGACATATGAATATAAAATACAGCAAAACTGTTAAAATATATTAATGGAACATTTTTTTACTTTGCATT
TTATATTGTTATTCACTTCTTATTTTTTTTTAAAAAAAAAAGCCTGAACAGTAAATTCAAAAGGAAAAGT
AATGATAATTAATTGTTGAGCATGGACCCAACTTGAAAAAAAAAAATGATGATGATAAATCTATAATCCTA
AAACCCTAAGTAAACACTTAAAAGATGTTCTGAAATCAGGAAAAGAATTATAGTATACTTTTGTGTTTCT
CTTTTATCAGTTGAAAAAAGGCACAGTAGCTCATGCCTGTAAGAACAGAGCTTTGGGAGTGCAAGGCAGG
CGGATCACTTGAGGCCAGGAGTTCCAGACCAGCCTGGGCAACATAGTGAAACCCCATCTCTACAAAAAAT
AAAAAAGAATTATTGGAATGTGTTTCTGTGTGCCTGTAATCCTAGCTATTCCGAAAGCTGAGGCAGGAGG
ATCTTTTGAGCCCAGGAGTTTGAGGTTACAGGGAGTTATGATGTGCCAGTGTACTCCAGCCTGGGGAACA
CCGAGACTCTGTCTTATTTAAAAAAAAAAAAAAAAAAAATGCTTGCAATAATGCCTGGCACATAGAAGGTA
ACAGTAAGTGTTAACTGTAATAACCCAGGTCTAAGTGTGTAAGGCAATAGAAAAATTGGGGCAAATAAGC
CTGACCTATGTATCTACAGAATCAGTTTGAGCTTAGGTAACAGACCTGTGGAGCACCAGTAATTACACAG
TAAGTGTTAACCAAAAGCATAGAATAGGAATATCTTGTTCAAGGGACCCCCAGCCTTATACATCTCAAGG
TGCAGAAAGATGACTTAATATAGGACCCATTTTTTCCTAGTTCTCCAGAGTTTTTATTGGTTCTTGAGAA
AGTAGTAGGGGAATGTTTTAGAAAATGAATTGGTCCAACTGAAATTACATGTCAGTAAGTTTTTATATAT
TGGTAAATTTTAGTAGACATGTAGAAGTTTTCTAATTAATCTGTGCCTTGAAACATTTTCTTTTTTCCTA
AAGTGCTTAGTATTTTTTCCGTTTTTTGATTGGTTACTTGGGAGCTTTTTTGAGGAAATTTAGTGAACTG
CAGAATGGGTTTGCAACCATTTGGTATTTTTGTTTTGTTTTTTAGAGGATGTATGTGTATTTTAACATTT
CTTAATCATTTTTAGCCAGCTATGTTTGTTTTGCTGATTTGACAAACTACAGTTAGACAGCTATTCTCAT
TTTGCTGATCATGACAAAATAATATCCTGAATTTTTAAATTTTGCATCCAGCTCTAAATTTTCTAAACAT
AAAATTGTCCAAAAAATAGTATTTTCAGCCACTAGATTGTGTGTTAAGTCTATTGTCACAGAGTCATTTT
ACTTTTAAGTATATGTTTTTACATGTTAATTATGTTTGTTATTTTTAATTTTAACTTTTTAAAATAATTC
CAGTCACTGCCAATACATGAAAAATTGGTCACTGGAATTTTTTTTTTGACTTTTATTTTAGGTTCATGTG
TACATGTGCAGGTGTGTTATACAGGTAAATTGCGTGTCATGAGGGTTTGGTGTACAGGTGATTTCATTAC
CCAGGTAATAAGCATAGTACCCAATAGGTAGTTTTTTGATCCTCACCCTTCTCCCACCCTCAAGTAGGCC
CTGGTGTTGCTGTTTCCTTCTTTGTGTCCATGTATACTCAGTGTTTAGCTCCCACTTAGAAGTGAGAACA
TGCGGTAGTTGGTTTTCTGTTCCTGGATTAGTTCACTTAGGATAATGACCTCTAGCTCCATCTGGTTTTT
ATGGCTGCATAGTATTCCATGGTGTATATGTATCACATTTTCTTTATCCAGTCTACCATTGATAGGCATT
TAGGTTGATTCCCTGTCTTTGTTATCATGAATAGTGCTGTGATGAACATACACATGCATGTGTCTTTATG
GTAGAAAAATTTGTATTCCTTTAGGTACATATAGAATAATGGGGTTGCTAGGGTGAATGGTAGTTCTATT
TTCAGTTATTTGAGAAATCTTCAAACTGCTTTTCATAATAGCTAAACTAATTTACAGTCCCGCCAGCAGT
GTATAAGTGTTCCCTTTTCTCCACAACCTTGCCAACATCTGTGATTTTTTGACTTTTTAATAATAGCCAT
TCCTAGAGAATTGATTTGCAATTCTCTATTAGTGATATTAAGCATTTTTTCATATGCTTTTTAGCTGTCT
GTATATATTCTTCTGAAAAATTTTCATGTCCTTTGCCCAGTTTGTAGTGGGGTGGGTTGTTTTTTGCTTG
TTAATTAGTTTTAAGTTCCTTCCAGATTCTGCATATCCCTTTGTTGGATACATGGTTTGCAGATATTTTT
CTCCCATTGTGTAGGTTGTCTTTTACTCTGTTGATAGTTTCTTTTGCCATGCAGGAGCTCGTTAGGTCCC
ATTTGTGTTTGTTTTTGTTGCAGTTGCTTTTGGCGTCTTCATCATAAAATCTGTGCCAGGGCCTATGTCC
AGAATGGTATTTCCTAGGTTGTCTTCCAGGGTTTTTACAATTTTAGATTTTACGTTTATGTCTTTAATCC
ATCTTGAGTTGATTTTTGTATATGGCACAAGGAAGGGGTCCAGTTTCACTCCAATTCCTATGGCTAGCAA
TTATCCCAGCACCATTTATTGAATACGGAGTCCTTTCCCCATTGCTTGTTTTTTGTCAACTTTGTTGAAG
ATCAGATGGTTGTAAGTGTGTGGCTTTATTTCTTGGCTCTCTATTCTCCATTGGTCTATGTGTCTGTTTT
TATAACAGTACCCTGCTGTTCAGGTTCCTATAGCCTTTTAGTATAAAATCGGCTAATGTGATGCCTCCAG
CTTTGTTCTTTTTGCTTAGGATTGCTTTGGCTATTTGGGCTCCTTTTTGGGTCCATATTAATTTTAAAAC
AGTTTTTTCTGGTTTTGTGAAGGATATCATTGGTAGTTTATAGGAATAGCATTGAATCTGTAGATTGCTT
TGGGCAGTATGGCCATTTTAACAATATTAATTCTTCCTATCTATGAATATGGAATGTTTTTCCATGTGTT
TGTGTCATCTCTTTATACCTGATGTATAAAGAAAAGCTGGTATTATTCCTACTCAATCTGTTCCAAAAAA
TTGAGGAGGAGGAACTCTTCCCTAATGAGGCCAGCATCATTCTGATACCAAAACCTGGCAGAGACACAAC
AGAAAAAAGAAAACTTCAGGCCAATATCCTTGATGAATATAGATGCAAAAATCCTCAACAAAATACTAGC
AAACCAAATCCAGCAGCACATCAAAAAGCTGATCTACTTTGATCAAGTAGGCTTTATCCCTGGGATGCAA
GGTTGGTTCAACATACACAAATCAATAAGTGTGATTCATCACATAAACAGAGCTAAAAACAAAAACCACA
AGATTATCTCAATAGGTAGAGAAAAGGTTGTCAATAAAATTTAACATCCTCCATGTTAAAAACCTTCAGT
AGGTCAGGTGTAGTGACTCACACCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGGCATATCTCTTAAG
CCCAGGAGTTCAAGACGAGCCTAGGCAGCATGGTGAAACCCCATCTCTACAAAAAAAAAAAAAAAAAAAA
ATTAGCTTGGTATGGTGACATGCACCTATAGTCCCAGCTATTCAGGAGGTTGAGGTGGGAGGATTGTTTG
AGCCCGGGAGGCAGAGGTTGGCAGCGAGCTGAGATCATGCCACCGCACTCCAGCCTGGGCAACGGAGTGA
GACCCTGTCTCAAAAAAGAAAAATCACAAACAATCCTAAACAAACTAGGCATTGAAGGAACATGCCTCAA
AAAAATAAGAACCATCTATGACAGACCCATAGCCAATATCTTACCAAATGGGCAAAAGCTGGAAGTATTC
TCCTTGAGAACCGTAACAAGACAAGGATGTCCACTCTCACCACTCCTTTTCAGCATAGTTCTGGAAGTCC
TAGCCAGAGCAATCAGGAAAGAGAAAGAAAGAAAGACATTCAGATAGGAAGAGAAGAAGTCAAACTATTT
CTGTTTGCAGGCAGTATAATTCTGTACCTAGAAAATCTCATAGTCTCTGCCCAGAAACTCCTAAATCTGT
TAAAAATTTCAGCAAAGTTTTGGCATTCTCTATACTCCAACACCTTCCAAAGTGAGAGCAAAATCAAGAA
CACAGTCCCATTCACAATAGCCGCAAAACGAATAAAATACCTAGGAATCCAGCTAACCAGGGAGGTGAAA
GATCTCTATGAGAATTACAAAACACTGCTGAAAGAAATCAGAGATGACACAAACAAATGGAAATGTTCTT
TTTTAACACCTTGCTTTATCTAATTCACTTATGATGAAGATACTCATTCAGTGGAACAGGTATAATAAGT
CCACTCGATTAAATATAAGCCTTATTCTCTTTCCAGAGCCCAAGAAGGGGCACTATCAGTGCCCAGTCAA
TAATGACGAAATGCTAATATTTTTCCCCTTTACGGTTTCTTTCTTCTGTAGTGTGGTACACTCGTTTCTT

[1134]

AAGATAAGGAAACTTGAACTACCTTCCTGTTTGCTTCTACACATACCCATTCTCTTTTTTTGCCACTCTG
GTCAGGTATAGGATGATCCCTACCACTTTCAGTTAAAAACTCCTCCTCTTACTAAATGTTCTCTTACCCT
CTGGCCTGAGTAGAACCTAGGGAAAATGGAAGAGAAAAAGATGAAAGGGAGGTGGGGCCTGGGAAGGGAA
TAAGTAGTCCTGTTTGTTTGTGTGTTTGCTTTAGCACCTGCTATATCCTAGGTGCTGTGTTAGGCACACA
TTATTTTAAGTGGCCATTATATTACTACTACTCACTCTGGTCGTTGCCAAGGTAGGTAGTACTTTCTTGG
ATAGTTGGTTCATGTTACTTACAGATGGTGGGCTTGTTGAGGCAAACCCAGTGGATAATCATCGGAGTGT
GTTCTCTAATCTCACTCAAATTTTTCTTCACATTTTTTGGTTTGTTTTGGTTTTTGATGGTAGTGGCTTA
TTTTTGTTGCTGGTTTGTTTTTTGTTTTTTTTTGAGATGGCAAGAATTGGTAGTTTTATTTATTAATTGC
CTAAGGGTCTCTACTTTTTTTAAAAGATGAGAGTAGTAAAATAGATTGATAGATACATACATACCCTTAC
TGGGGACTGCTTATATTCTTTAGAGAAAAAATTACATATTAGCCTGACAAACACCAGTAAAATGTAAATA
TATCCTTGAGTAAATAAATGAATGTATATTTTGTGTCTCCAAATATATATCTATATTCTTACAAATGT
GTTTATATGTAATATCAATTTATAAGAACTTAAAATGTTGGCTCAAGTGAGGGATTGTGGAAGGTAGCAT
TATATGGCCATTTCAACATTTGAACTTTTTTCTTTTCTTCATTTTCTTCTTTTCTTCAGGAATATTTTTC
AAGATGTCTTACACAGAGACACTCTAGTGAAAGCCTTCCTGGATCAGGTAAATGTTGAACTTGAGATTGT
CAGAGTGAATGATATGACATGTTTTCTTTTTTAATATATCCTACAATGCCTGTTCTATATATTTATATTC
CCCTGGATCATGCCCCAGAGTTCTGCTCAGCAATTGCAGTTAAGTTAGTTACACTACAGTTCTCAGAAGA
GTCTGTGAGGGCATGTCAAGTGCATCATTACATTGGTTGCCTCTTGTCCTAGATTTATGCTTCGGGAATT
CAGACCTTTGTTTACAATATAATAAATATTATTGCTATCTTTTAAAGATATAATAATAAGATATAAAGTT
GACCACAACTACTGTTTTTTGAAACATAGAATTCCTGGTTTACATGTATCAAAGTGAAATCTGACTTAGC
TTTTACAGATATAATATATACATATATATATCCTGCAATGCTTGTACTATATATGTAGTACAAGTATATA
TATATGTTTGTGTGTGTATATATATATAGTACGAGCATATATACATATTACCAGCATTGTAGGATATATA
TATGTTTATATATTAAAAAAAAGTTATAAACTTAAAACCCTATTATGTTATGTAGAGTATATGTTATATA
TGATATGTAAAATATATAACATATACTCTATGATAGAGTGTAATATATTTTTTATATATATTTTAACATT
TATAAAATGATAGAATTAAGAATTGAGTCCTAATCTGTTTTATTAGGTGCTTTTTGTAGTGTCTGGTCTT
TCTAAAGTGTCTAAATGATTTTTCCTTTTGACTTATTAATGGGGAAGAGCCTGTATATTAACAATTAAGA
GTGCAGCATTCCATACGTCAAACAACAAACATTTTAATTCAAGCATTAACCTATAACAAGTAAGTTTTTT
TTTTTTTTTTGAGAAAGGGAGGTTGTTTATTTGCCTGAAATGACTCAAAAATATTTTTGAAACATAGTGT
ACTTATTTAAATAACATCTTTATTGTTTCATTCTTTTAAAAAATATCTACTTAATTACACAGTTGAAGGA
AATCGTAGATTATATGGAACTTATTTCTTAATATATTACAGTTTGTTATAATAACATTCTGGGGATCAGG
CCAGGAAACTGTGTCATAGATAAAGCTTTGAAATAATGAGATCCTTATGTTTACTAGAAATTTTGGATTG
AGATCTATGAGGTCTGTGACATATTGCGAAGTTCAAGGAAAATTCGTAGGCCTGGAATTTCATGCTTCTC
AAGCTGACATAAAATCCCTCCCACTCTCCACCTCATCATATGCACACATTCTACTCCTACCCACCCACTC
[1135] CACCCCCTGCAAAAGTACAGGTATATGAATGTCTCAAAACCATAGGCTCATCTTCTAGGAGCTTCAATGT
TATTTGAAGATTTGGGCAGAAAAAATTAAGTAATACGAAATAACTTATGTATGAGTTTTAAAAGTGAAGT
AAACATGGATGTATTCTGAAGTAGAATGCAAAATTTGAATGCATTTTTAAAGATAAATTAGAAAACTTCT
AAAAACTGTCAGATTGTCTGGGCCTGGTGGCTTATGCCTGTAATCCCAGCACTTTGGGAGTCCGAGGTGG
GTGGATCACAAGGTCAGGAGATCGAGACCATCCTGCCAACATGGTGAAACCCCGTCTCTACTAAGTATAC
AAAAATTAGCTGGGCGTGGCAGCGTGTGCCTGTAATCCCAGCTACCTGGGAGGCTGAGGCAGGAGAATCG
CTTGAACCCAGGAGGTGTAGGTTGCAGTGAGTCAAGATCGCGCCACTGCACTTTAGCCTGGTGACAGAGC
TAGACTCCGTCTCAAAAAAAAAAAAAAAATATCAGATTGTTCCTACACCTAGTGCTTCTATACCACACTCC
TGTTAGGGGGCATCAGTGGAAATGGTTAAGGAGATGTTTAGTGTGTATTGTCTGCCAAGCACTGTCAACA
CTGTCATAGAAACTTCTGTACGAGTAGAATGTGAGCAAATTATGTGTTGAAATGGTTCCTCTCCCTGCAG
GTCTTTCAGCTGAAACCTGGCTTATCTCTCAGAAGTACTTTCCTTGCACAGTTTCTACTTGTCCTTCACA
GAAAAGCCTTGACACTAATAAAATATATAGAAGACGATACGTGAGTAAAACTCCTACACGGAAGAAAAAC
CTTTGTACATTGTTTTTTTGTTTTGTTTCCTTTGTACATTTTCTATATCATAATTTTTGCGCTTCTTTTT
TTTTTTTTTTTTTTTTTTTTTCCATTATTTTTAGGCAGAAGGGAAAAAAGCCCTTTAAATCTCTTCGGAA
CCTGAAGATAGACCTTGATTTAACAGCAGAGGGCGATCTTAACATAATAATGGCTCTGGCTGAGAAAATT
AAACCAGGCCTACACTCTTTTATCTTTGGAAGACCTTTCTACACTAGTGTGCAAGAACGAGATGTTCTAA
TGACTTTTTAAATGTGTAACTTAATAAGCCTATTCCATCACAATCATGATCGCTGGTAAAGTAGCTCAGT
GGTGTGGGGAAACGTTCCCCTGGATCATACTCCAGAATTCTGCTCTCAGCAATTGCAGTTAAGTAAGTTA
CACTACAGTTCTCACAAGAGCCTGTGAGGGGATGTCAGGTGCATCATTACATTGGGTGTCTCTTTTCCTA
GATTTATGCTTTTGGGATACAGACCTATGTTTACAATATAATAAATATTATTGCTATCTTTTAAAGATAT
AATAATAGGATGTAAACTTGACCACAACTACTGTTTTTTTGAAATACATGATTCATGGTTTACATGTGTC
AAGGTGAAATCTGAGTTGGCTTTTACAGATAGTTGACTTTCTATCTTTTGGCATTCTTTGGTGTGTAGAA
TTACTGTAATACTTCTGCAATCAACTGAAAACTAGAGCCTTTAAATGATTTCAATTCCACAGAAAGAAAG
TGAGCTTGAACATAGGATGAGCTTTAGAAAGAAAATTGATCAAGCAGATGTTTAATTGGAATTGATTATT
AGATCCTACTTTGTGGATTTAGTCCCTGGGATTCAGTCTGTAGAAATGTCTAATAGTTCTCTATAGTCCT
TGTTCCTGGTGAACCACAGTTAGGGTGTTTTGTTTATTTTATTGTTCTTGCTATTGTTGATATTCTATGT
AGTTGAGCTCTGTAAAAGGAAATTGTATTTTATGTTTTAGTAATTGTTGCCAACTTTTTAAATTAATTTT
CATTATTTTTGAGCCAAATTGAAATGTGCACCTCCTGTGCCTTTTTTCTCCTTAGAAAATCTAATTACTT
GGAACAAGTTCAGATTTCACTGGTCAGTCATTTTCATCTTGTTTTCTTCTTGCTAAGTCTTACCATGTAC
CTGCTTTGGCAATCATTGCAACTCTGAGATTATAAAATGCCTTAGAGAATATACTAACTAATAAGATCTT
TTTTTCAGAAACAGAAAATAGTTCCTTGAGTACTTCCTTCTTGCATTTCTGCCTATGTTTTTGAAGTTGT
TGCTGTTTGCCTGCAATAGGCTATAAGGAATAGCAGGAGAAATTTTACTGAAGTGCTGTTTTCCTAGGTG
CTACTTTGGCAGAGCTAAGTTATCTTTTGTTTTCTTAATGCGTTTGGACCATTTTGCTGGCTATAAAATA
ACTGATTAATATAATTCTAACACAATGTTGACATTGTAGTTACACAAACACAAATAAATATTTTATTTAA

AATTCTGGAAGTAATATAAAAGGGAAAATATATTTATAAGAAAGGGATAAAGGTAATAGAGCCCTTCTGC
CCCCCACCCACCAAATTTACACAACAAAATGACATGTTCGAATGTGAAAGGTCATAATAGCTTTCCCATC
ATGAATCAGAAAGATGTGGACAGCTTGATGTTTTAGACAACCACTGAACTAGATGACTGTTGTACTGTAG
CTCAGTCATTTAAAAAATATATAAATACTACCTTGTAGTGTCCCATACTGTGTTTTTTACATGGTAGATT
CTTATTTAAGTGCTAACTGGTTATTTTCTTTGGCTGGTTTATTGTACTGTTATACAGAATGTAAGTTGTA
CAGTGAAATAAGTTATTAAAGCATGTGTAAACATTGTTATATATCTTTTCTCCTAAATGGAGAATTTTGA
ATAAAATATATTTGAAATTTT

SEQ ID NO:134

SEQ ID NO:133 SEQ ID NO:13 的反向互补序列

AAAATTTCAAATATATTTTATTCAAAATTCTCCATTTAGGAGAAAAGATATATAACAATGTTTACACATGCTTTAAT
AACTTATTTCACTGTACAACTTACATTCTGTATAACAGTACAATAAACCAGCCAAAGAAAATAACCAGTTAGCACTT
AAATAAGAATCTACCATGTAAAAAACACAGTATGGGACACTACAAGGTAGTATTTATATATTTTTTAAATGACTGAG
CTACAGTACAACAGTCATCTAGTTCAGTGGTTGTCTAAAACATCAAGCTGTCCACATCTTTCTGATTCATGATGGGA
AAGCTATTATGACCTTTCACATTCGAACATGTCATTTTGTTGTGTAAATTTGGTGGGTGGGGGGCAGAAGGGCTCTA
TTACCTTTATCCCTTTCTTATAAATATATTTTCCCTTTTATATTACTTCCAGAATTTTAAATAAAATATTTATTTGT
GTTTGTGTAACTACAATGTCAACATTGTGTTAGAATTATATTAATCAGTTATTTTATAGCCAGCAAAATGGTCCAAA
CGCATTAAGAAAACAAAAGATAACTTAGCTCTGCCAAAGTAGCACCTAGGAAAACAGCACTTCAGTAAAATTTCTCC
TGCTATTCCTTATAGCCTATTGCAGGCAAACAGCAACAACTTCAAAAACATAGGCAGAAATGCAAGAAGGAAGTACT
CAAGGAACTATTTTCTGTTTCTGAAAAAAAGATCTTATTAGTTAGTATATTCTCTAAGGCATTTTATAATCTCAGAG
TTGCAATGATTGCCAAAGCAGGTACATGGTAAGACTTAGCAAGAAGAAAACAAGATGAAAATGACTGACCAGTGAAA
TCTGAACTTGTTCCAAGTAATTAGATTTTCTAAGGAGAAAAAAGGCACAGGAGGTGCACATTTCAATTTGGCTCAAA
AATAATGAAAATTAATTTAAAAAGTTGGCAACAATTACTAAAACATAAAATACAATTTCCTTTTACAGAGCTCAACT
ACATAGAATATCAACAATAGCAAGAACAATAAAATAAACAAAACACCCTAACTGTGGTTCACCAGGAACAAGGACTA
TAGAGAACTATTAGACATTTCTACAGACTGAATCCCAGGGACTAAATCCACAAAGTAGGATCTAATAATCAATTCCA
ATTAAACATCTGCTTGATCAATTTTCTTTCTAAAGCTCATCCTATGTTCAAGCTCACTTTCTTTCTGTGGAATTGAA
ATCATTTAAAGGCTCTAGTTTTCAGTTGATTGCAGAAGTATTACAGTAATTCTACACACCAAAGAATGCCAAAAGAT
AGAAAGTCAACTATCTGTAAAAGCCAACTCAGATTTCACCTTGACACATGTAAACCATGAATCATGTATTTCAAAAA
AACAGTAGTTGTGGTCAAGTTTACATCCTATTATTATATCTTTAAAAGATAGCAATAATATTTATTATATTGTAAAC
ATAGGTCTGTATCCCAAAAGCATAAATCTAGGAAAAGAGACACCCAATGTAATGATGCACCTGACATCCCCTCACAG
GCTCTTGTGAGAACTGTAGTGTAACTTACTTAACTGCAATTGCTGAGAGCAGAATTCTGGAGTATGATCCAGGGGAA
CGTTTCCCCACACCACTGAGCTACTTTACCAGCGATCATGATTGTGATGGAATAGGCTTATTAAGTTACACATTTAA
AAAGTCATTAGAACATCTCGTTCTTGCACACTAGTGTAGAAAGGTCTTCCAAAGATAAAAGAGTGTAGGCCTGGTTT
AATTTTCTCAGCCAGAGCCATTATTATGTTAAGATCGCCCTCTGCTGTTAAATCAAGGTCTATCTTCAGGTTCCGAA
GAGATTTAAAGGGCTTTTTTCCCTTCTGCCTAAAAATAATGGAAAAAAAAAAAAAAAAAAAAAAAAAGAAGCGCAA
AAATTATGATATAGAAAATGTACAAAGGAAACAAAACAAAAAAACAATGTACAAAGGTTTTTCTTCCGTGTAGGAGT
TTTACTCACGTATCGTCTTCTATATATTTTATTAGTGTCAAGGCTTTTCTGTGAAGGACAAGTAGAAACTGTGCAAG
GAAAGTACTTCTGAGAGATAAGCCAGGTTTCAGCTGAAAGACCTGCAGGGAGAGGAACCATTTCAACACATAATTTG
CTCACATTCTACTCGTACAGAAGTTTCTATGACAGTGTTGACAGTGCTTGGCAGACAATACACACTAAACATCTCCT
TAACCATTTCCACTGATGCCCCCTAACAGGAGTGTGGTATAGAAGCACTAGGTGTAGGAACAATCTGATATTTTTTT
TTTTTTTTGAGACGGAGTCTAGCTCTGTCACCAGGCTAAAGTGCAGTGGCGCGATCTTGACTCACTGCAACCTACAC
CTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCAGGTAGCTGGGATTACAGGCACACGCTGCCACGCCCAGCT
AATTTTTGTATACTTAGTAGAGACGGGGTTTCACCATGTTGGCAGGATGGTCTCGATCTCCTGACCTTGTGATCCAC
CCACCTCGGACTCCCAAAGTGCTGGGATTACAGGCATAAGCCACCAGGCCCAGACAATCTGACAGTTTTTAGAAGTT
TTCTAATTTATCTTTAAAAATGCATTCAAATTTTGCATTCTACTTCAGAATACATCCATGTTTACTTCACTTTTAAA
ACTCATACATAAGTTATTTCGTATTACTTAATTTTTTCTGCCCAAATCTTCAAATAACATTGAAGCTCCTAGAAGAT
GAGCCTATGGTTTTGAGACATTCATATACCTGTACTTTTGCAGGGGGTGGAGTGGGTGGGTAGGAGTAGAATGTGTG
CATATGATGAGGTGGAGAGTGGGAGGGATTTTATGTCAGCTTGAGAAGCATGAAATTCCAGGCCTACGAATTTTCCT
TGAACTTCGCAATATGTCACAGACCTCATAGATCTCAATCCAAAATTTCTAGTAAACATAAGGATCTCATTATTTCA
AAGCTTTATCTATGACACAGTTTCCTGGCCTGATCCCCAGAATGTTATTATAACAAACTGTAATATATTAAGAAATA
AGTTCCATATAATCTACGATTTCCTTCAACTGTGTAATTAAGTAGATATTTTTTAAAAGAATGAAACAATAAAGATG
TTATTTAAATAAGTACACTATGTTTCAAAAATATTTTTGAGTCATTTCAGGCAAATAAACAACCTCCCTTTCTCAAA
AAAAAAAAAAAAACTTACTTGTTATAGGTTAATGCTTGAATTAAAATGTTTGTTGTTTGACGTATGGAATGCTGCAC
TCTTAATTGTTAATATACAGGCTCTTCCCCATTAATAAGTCAAAAGGAAAAATCATTTAGACACTTTAGAAAGACCA
GACACTACAAAAAGCACCTAATAAAACAGATTAGGACTCAATTCTTAATTCTATCATTTTATAAATGTTAAAATATA
TATAAAAAATATATTACACTCTATCATAGAGTATATGTTATATATTTTACATATCATATATAACATATACTCTACAT
AACATAATAGGGTTTTAAGTTTATAACTTTTTTTTAATATATAAACATATATATATCCTACAATGCTGGTAATATGT
ATATATGCTCGTACTATATATATATACACACACAAACATATATATATACTTGTACTACATATATAGTACAAGCATTG
CAGGATATATATATGTATATATTATATCTGTAAAAGCTAAGTCAGATTTCACTTTGATACATGTAAACCAGGAATTC
TATGTTTCAAAAAACAGTAGTTGTGGTCAACTTTATATCTTATTATTATATCTTTAAAAGATAGCAATAATATTTAT
TATATTGTAAACAAAGGTCTGAATTCCCGAAGCATAAATCTAGGACAAGAGGCAACCAATGTAATGATGCACTTGAC
ATGCCCTCACAGACTCTTCTGAGAACTGTAGTGTAACTAACTTAACTGCAATTGCTGAGCAGAACTCTGGGGCATGA
TCCAGGGGAATATAAATATATAGAACAGGCATTGTAGGATATATTAAAAAAGAAAACATGTCATATCATTCACTCTG
ACAATCTCAAGTTCAACATTTACCTGATCCAGGAAGGCTTTCACTAGAGTGTCTCTGTGTAAGACATCTTGAAAAAT
ATTCCTGAAGAAAAGAAGAAAATGAAGAAAAGAAAAAAGTTCAAATGTTGAAATGGCCATATAATGCTACCTTCCAC

[1136]

AATCCCTCACTTGAGCCAACATTTTAAGTTCTTATAAATTGATATTACATATAAACACATTTGTAAGAATATAGATA
TATATATTTGGAGACACAAAATATACATTCATTTATTTACTCAAGGATATATTTACATTTTACTGGTGTTTGTCAGG
CTAATATGTAATTTTTTCTCTAAAGAATATAAGCAGTCCCCAGTAAGGGTATGTATGTATCTATCAATCTATTTTAC
TACTCTCATCTTTTAAAAAAAGTAGAGACCCTTAGGCAATTAATAAATAAAACTACCAATTCTTGCCATCTCAAAAA
AAAACAAAAAACAAACCAGCAACAAAAATAAGCCACTACCATCAAAAACCAAAACAAACCAAAAAATGTGAAGAAAA
ATTTGAGTGAGATTAGAGAACACACTCCGATGATTATCCACTGGGTTTGCCTCAACAAGCCCACCATCTGTAAGTAA
CATGAACCAACTATCCAAGAAAGTACTACCTACCTTGGCAACGACCAGAGTGAGTAGTAGTAATATAATGGCCACTT
AAAATAATGTGTGCCTAACACAGCACCTAGGATATAGCAGGTGCTAAAGCAAACACACAAACAAACAGGACTACTTA
TTCCCTTCCCAGGCCCCACCTCCCTTTCATCTTTTTCTCTTCCATTTTCCCTAGGTTCTACTCAGGCCAGAGGGTAA
GAGAACATTTAGTAAGAGGAGGAGTTTTTAACTGAAAGTGGTAGGGATCATCCTATACCTGACCAGAGTGGCAAAAA
AAGAGAATGGGTATGTGTAGAAGCAAACAGGAAGGTAGTTCAAGTTTCCTTATCTTAAGAAACGAGTGTACCACACT
ACAGAAGAAAGAAACCGTAAAGGGGAAAAATATTAGCATTTCGTCATTATTGACTGGGCACTGATAGTGCCCCTTCT
TGGGCTCTGGAAAGAGAATAAGGCTTATATTTAATCGAGTGGACTTATTATACCTGTTCCACTGAATGAGTATCTTC
ATCATAAGTGAATTAGATAAAGCAAGGTGTTAAAAAAGAACATTTCCATTTGTTTGTGTCATCTCTGATTTCTTTCA
GCAGTGTTTTGTAATTCTCATAGAGATCTTTCACCTCCCTGGTTAGCTGGATTCCTAGGTATTTTATTCGTTTTGCG
GCTATTGTGAATGGGACTGTGTTCTTGATTTTGCTCTCACTTTGGAAGGTGTTGGAGTATAGAGAATGCCAAAACTT
TGCTGAAATTTTTAACAGATTTAGGAGTTTTCTGGGCAGAGACTATGAGATTTTCTAGGTACAGAATTATACTGCCTG
CAAACAGAAATAGTTTGACTTCTTCTCTTCCTATCTGAATGTCTTTCTTTCTTTCTCTTTCCTGATTGCTCTGGCTA
GGACTTCCAGAACTATGCTGAAAAGGAGTGGTGAGAGTGGACATCCTTGTCTTGTTACGGTTCTCAAGGAGAATACT
TCCAGCTTTTGCCCATTTGGTAAGATATTGGCTATGGGTCTGTCATAGATGGTTCTTATTTTTTTGAGGCATGTTCC
TTCAATGCCTAGTTTGTTTAGGATTGTTTGTGATTTTTCTTTTTTGAGACAGGGTCTCACTCCGTTGCCCAGGCTGG
AGTGCGGTGGCATGATCTCAGCTCGCTGCCAACCTCTGCCTCCCGGGCTCAAACAATCCTCCCACCTCAACCTCCTG
AATAGCTGGGACTATAGGTGCATGTCACCATACCAAGCTAATTTTTTTTTTTTTTTTTTTTGTAGAGATGGGGTTT
CACCATGCTGCCTAGGCTCGTCTTGAACTCCTGGGCTTAAGAGATATGCCCGCCTTGGCCTCCCAAAGTGCTGGGAT
TACAGGTGTGAGTCACTACACCTGACCTACTGAAGGTTTTTAACATGGAGGATGTTAAATTTTATTGACAACCTTTT
CTCTACCTATTGAGATAATCTTGTGGTTTTTGTTTTTAGCTCTGTTTATGTGATGAATCACACTTATTGATTTGTGT
ATGTTGAACCAACCTTGCATCCCAGGGATAAAGCCTACTTGATCAAAGTAGATCAGCTTTTTGATGTGCTGCTGGAT
TTGGTTTGCTAGTATTTTGTTGAGGATTTTTGCATCTATATTCATCAAGGATATTGGCCTGAAGTTTTCTTTTTTCT
GTTGTGTCTCTGCCAGGTTTTGGTATCAGAATGATGCTGGCCTCATTAGGGAAGAGTTCCTCCTCCTCAATTTTTTG
GAACAGATTGAGTAGGAATAATACCAGCTTTTCTTTATACATCAGGTATAAAGAGATGACACAAACACATGGAAAAA
CATTCCATATTCATAGATAGGAAGAATTAATATTGTTAAAATGGCCATACTGCCCAAAGCAATCTACAGATTCAATG
CTATTCCTATAAACTACCAATGATATCCTTCACAAAACCAGAAAAAACTGTTTTAAAATTAATATGGACCCAAAAAG
GAGCCCAAATAGCCAAAGCAATCCTAAGCAAAAAGAACAAAGCTGGAGGCATCACATTAGCCGATTTTATACTAAAA
[1137] GGCTATAGGAACCTGAACAGCAGGGTACTGTTATAAAAACAGACACATAGACCAATGGAGAATAGAGAGCCAAGAAA
TAAAGCCACACACTTACAACCATCTGATCTTCAACAAAGTTGACAAAAAACAAGCAATGGGGAAAGGACTCCGTATT
CAATAAATGGTGCTGGGATAATTGCTAGCCATAGGAATTGGAGTGAAACTGGACCCCTTCCTTGTGCCATATACAAA
AATCAACTCAAGATGGATTAAAGACATAAACGTAAAATCTAAAATTGTAAAAACCCTGGAAGACAACCTAGGAAATA
CCATTCTGGACATAGGCCCTGGCACAGATTTTATGATGAAGACGCCAAAAGCAACTGCAACAAAAACAAACACAAAT
GGGACCTAACGAGCTCCTGCATGGCAAAAGAAACTATCAACAGAGTAAAAGACAACCTACACAATGGGAGAAAAATA
TCTGCAAACCATGTATCCAACAAAGGGATATGCAGAATCTGGAAGGAACTTAAAACTAATTAACAAGCAAAAAACAA
CCCACCCCACTACAAACTGGGCAAAGGACATGAAAATTTTTCAGAAGAATATATACAGACAGCTAAAAAGCATATGA
AAAAATGCTTAATATCACTAATAGAGAATTGCAAATCAATTCTCTAGGAATGGCTATTATTAAAAAGTCAAAAAATC
ACAGATGTTGGCAAGGTTGTGGAGAAAAGGGAACACTTATACACTGCTGGCGGGACTGTAAATTAGTTTAGCTATTA
TGAAAAGCAGTTTGAAGATTTCTCAAATAACTGAAAATAGAACTACCATTCACCCTAGCAACCCCATTATTCTATAT
GTACCTAAAGGAATACAAATTTTTCTACCATAAAGACACATGCATGTGTATGTTCATCACAGCACTATTCATGATAA
CAAAGACAGGGAATCAACCTAAATGCCTATCAATGGTAGACTGGATAAAGAAAATGTGATACATATACACCATGGAA
TACTATGCAGCCATAAAAACCAGATGGAGCTAGAGGTCATTATCCTAAGTGAACTAATCCAGGAACAGAAAACCAAC
TACCGCATGTTCTCACTTCTAAGTGGGAGCTAAACACTGAGTATACATGGACACAAAGAAGGAAACAGCAACACCAG
GGCCTACTTGAGGGTGGGAGAAGGGTGAGGATCAAAAAACTACCTATTGGGTACTATGCTTATTACCTGGGTAATGA
AATCACCTGTACACCAAACCCTCATGACACGCAATTTACCTGTATAACACACCTGCACATGTACACATGAACCTAAA
ATAAAAGTCAAAAAAAAAATTCCAGTGACCAATTTTTCATGTATTGGCAGTGACTGGAATTATTTTAAAAAGTTAAA
ATTAAAAATAACAAACATAATTAACATGTAAAAACATATACTTAAAAGTAAAATGACTCTGTGACAATAGACTTAAC
ACACAATCTAGTGGCTGAAAATACTATTTTTTGGACAATTTTATGTTTAGAAAATTTAGAGCTGGATGCAAAATTTA
AAAATTCAGGATATTATTTTGTCATGATCAGCAAAATGAGAATAGCTGTCTAACTGTAGTTTGTCAAATCAGCAAAA
CAAACATAGCTGGCTAAAAATGATTAAGAAATGTTAAAATACACATACATCCTCTAAAAAACAAAACAAAAATACCA
AATGGTTGCAAACCCATTCTGCAGTTCACTAAATTTCCTCAAAAAAGCTCCCAAGTAACCAATCAAAAAACGGAAAA
AATACTAAGCACTTTAGGAAAAAAGAAAATGTTTCAAGGCACAGATTAATTAGAAAACTTCTACATGTCTACTAAAA
TTTACCAATATATAAAAACTTACTGACATGTAATTTCAGTTGGACCAATTCATTTTCTAAAACATTCCCCTACTACT
TTCTCAAGAACCAATAAAAACTCTGGAGAACTAGGAAAAAATGGGTCCTATATTAAGTCATCTTTCTGCACCTTGAG
ATGTATAAGGCTGGGGGTCCCTTGAACAAGATATTCCTATTCTATGCTTTTGGTTAACACTTACTGTGTAATTACTG
GTGCTCCACAGGTCTGTTACCTAAGCTCAAACTGATTCTGTAGATACATAGGTCAGGCTTATTTGCCCCAATTTTTC
TATTGCCTTACACACTTAGACCTGGGTTATTACAGTTAACACTTACTGTTACCTTCTATGTGCCAGGCATTATTGCA
AGCATTTTTTTTTTTTTTTTTTTAAATAAGACAGAGTCTCGGTGTTCCCCAGGCTGGAGTACACTGGCACATCATAA
CTCCCTGTAACCTCAAACTCCTGGGCTCAAAAGATCCTCCTGCCTCAGCTTTCGGAATAGCTAGGATTACAGGCACA
CAGAAACACATTCCAATAATTCTTTTTTTATTTTTTGTAGAGATGGGGTTTCACTATGTTGCCCAGGCTGGTCTGGAA
CTCCTGGCCTCAAGTGATCCGCCTGCCTTGCACTCCCAAAGCTCTGTTCTTACAGGCATGAGCTACTGTGCCTTTTT

TCAACTGATAAAAGAGAAACACAAAAGTATACTATAATTCTTTTCCTGATTTCAGAACATCTTTTAAGTGTTTACTT
AGGGTTTTAGGATTATAGATTTATCATCATCATTTTTTTTTTCAAGTTGGGTCCATGCTCAACAATTAATTATCATT
ACTTTTCCTTTTGAATTTACTGTTCAGGCTTTTTTTTTTAAAAAAAAATAAGAAGTGAATAACAATATAAAATGCAA
AGTAAAAAAATGTTCCATTAATATATTTTAACAGTTTTGCTGTATTTTATATTCATATGTCAAGAAAAAAAAGACTT
AGAATTACTTCTTTGGTTTTGTTAATAAGTCCATTCAATAACTGTTACATAATAGACTATACACAATGTAGGGGCTG
GGGCTGGGGCTGTCTGGTTGGCCACTATTAAGCCCGGTACCTAGAACAGTGCCTAGGAGCTCAATATATATTTATTA
AACAAATAGTTAAATATTTACTTTAAGGGTTAATCTTACCTATTTTAGGGTAAACACAATTGTATGGAAAAAGAAGA
ACATTTAAGAAAAAAATGCTTGATTAAATGTTTCTTCAAGCATAATTCTGCTGGTAATATTTCTTTCTTTTTAAATG
CAACTTCTGTTTTGTTTTTATTCGTAAAAGACACAATTTCATATTGCTTGACTACAGTACCAGCAGGCAGAGCATTA
CGTACAAATCAGGAGTAAAGCTTTCGTCAGTGTAGATGATCGTATCCTGAGCCATGTCTTCTTCTGAAGTGGCTCTC
CAGAAGGCTGTCAGCTCGGATCTCATGTATCTACGCTGATTATAAATATGTTCATGACAGGGTGGCATCTGCTTCAC
AGTATTGACATCCACATCTATGTGTGTGGTGGGATATGGAGCATACATGACTTGCCGGAAAGGCAGCACAAAGCTTC
CAGTTGAATCCTGTCAAAATAAAAGGAAAATTTACTGTCTTACATGCCAAACGATATGAATAATTGTTTTTTAATTT
TAAAAAATGGTCTCTGATTTAGCAAATCCATCTCTAAAAATTTATCCTAAGAAGCTATTCATGAGTGTGTGCAAAGT
TTTAGTGACAGAAGTGTTCAGTATAGCATTGGTTTCAATAGCAAAACATCAGAATGACCTAAGCATCCAAAGGAAGT
GGTTAAGTAAATAAAAATGAATTTATTATTTGGCATCTATAGATGAAGTTGTTGAAGTTTGTGTTGATATTCAAAGA
GGCTTAGAATCTGCTGTGAAAAAAGCAGATTAGAACATGCAGAGATGATCCAGTGGGGGTGGGAGTCTTAAAGAAAA
TACATGCAAAAGGGAAAAAAAGACTAGAAAGAGATATACCAAATATTAACAGTGGGTATTCACTGGATTCAAAATTA
CAAGTGTCTTTTAATTTCTTAATTATTCTATGCACCTCAATGAGTATGTATTTTTATAAAACAAATTTACTTCTTTC
CTAATCAGTTGGGGTAGCCAGTTACCATGTAATAATAACCTACTGCGTATGAGGCACTGTGCTATGAACAGAACAGA
AGAAAGAGGCTATTTCCATTTTTTTACTATTATAATGCTGTGATAATCATCTTTGTATAATGTCAAAGGTATTTTCT
GGAAAACTATATGATTTGTATTTGCATGGAGAACTATTAAATATTTTTTCATTAGATCATAGTTTTGAAATTTCCTC
TAGGCATCTGGTATAGGTCCCTTTCAGAACTAAAATTATATAATTAAATGAGAGGGTGCTGAGGAAACAAAAGTATT
AAGCACTTTATTATGCCTTTAACAGGAGAAGTTCTAATTTACAGATATGTATAACTCAAATAGAAACTACGTTTCCT
TTAACTAGTTTCTATATTTAAATACTATCATTCTGAGTTCATAGCAGCACATTATTGAAAATGCACAAAAGCCTAAT
ATAACTTGTATTTGCAACAATTTTTAAATTTTTTTATTTCATTTTGCCGTGGCTACATCCTCAAGAAAGTAGTCACT
ATGAGACAATCACATAACCATTAGGAAAGCTATTTCTTCACGTCTTGGGACTATAATTGAAATAACAATAATAGATT
CAACAAGGTGAACAACTTTCTTCCCTTTAAGTATTATAAATAATTGCTAATTTACATTTCTCTGTTTTGCTCCTTGA
CAGTAGACTCAAATAAAAGAAAAATACAATTTTTTTCTAATATATATTATATATAGAAATATATATTTTAGTACA
CATACATAGGAATTGCTTAAATCCTATAAGCTTCTTAAAGATGTTTAAAATTTTTTTTCAATAAAATTCAAATCTAT
TTAGTTAAAAAGAAATGATCTTATTTTTGATGTGCAACCTGATTTAAGCATGTGTTAAAAAAAAAAAAGTCCTGAGG
CTAGACATGTAGGAACAGGGACCCACCTGGAACACAAAGGGTATTCTATGGTGTTTCACTGATGATACTAACTATAA
ATCCATAAGACATATAGTCTATGTGCAGAACTGTGTAAAGGAAGTCAGTCTCTGGGCATGTCAATATGAGATACATT
[1138] AAATGCTAATATTTAAGATTTGTCTATAAAGAGTCTCAAAAATGATTTTAGAAAAGTGGTTTCACTTGTGATAACTA
GAAACTATACCTTTAGCAGGCCTTGTACAAAGAGCCCTGACTCATATTTAAATGATGATTCTGCTTCACATAACCTG
GAGCATTTTCTCTCTGCTGGAGTCAGAAAAAGGCATAATGTTCTGACTATCTATAAAAGAAAATATTTTAGCATTAA
AACATGAAGTAAAAAGACCACTGATTTGCTTATGAAAGATATCTGAAATTTTAATTGTTATTATCAATAAAACATAT
CCTAAGAAATAAGTATTCTTTAGTCACCTGGAATCCATGGGATGTAACAGTTGTTGCTTAAAATAGGGATTCTGTGG
TCAAATAAATTTAGAAAATGCTTGGTAAATTTACTGTATGGCATCTCAAAACCTTTAACATTTGGCTACATACTGTG
ACTCTTCAGGTGAATTATATCATCTGCAGACTTTCTCAGACTTATTTGACCATGCAACTTTTATAGCTTCTGTTTAG
TGGGTCTACACATGAAATTCGTTGTAAGAAATATCAAAGAATGTCCAGATCCTCCAAAAAGAAGAGAATTAATTATT
GAATTTGATTTAAAATAAGCAGGTCATTGGGTGGGATTTAGAAATCTGATTCTAATAATATTTGACCTACAGCTCTA
TTAGGAAAAATAAAAAGGCATGTAACTATCTTGAAATTCAAACCATATCCACAGTTATTATGTTAAAGGAGTGGTTT
TCAACCTAGAGTGGTGGCAACTTCTTTCTCAAATTAAAACAAGATAAAAGGTAAGCTATTCTCCCTCAAGAGGATGG
AGGATGGGAAAAATGTTTAAAAAAAAAAAAAAAAACACTCCAACTATGGAGCCTTTCTCCCTTCATAAAGCAGCTCG
GCAGTCACTCTGTGCAACCTAAGGCTTTGGAGATCACAGATGGAAAGCCACCTGTTTGAGGTAACAGAAGGAATAAG
GTCACTAGTTCGTAGATGCAATATAATGACACAGGTATACTAAGCTCTCATAAATGGTTATATGAGAAATATAAATT
AAGGCTCATGTAAATATACAAAGTAGCTGATTACAAAAAAAAATTATGAATATCTTTGTAAAGTATCATTTCCAACA
TATTTCCTATGTAAAACTTTTTTTAAAAAATTAGGTTTGCTGAAATTGAAAGATACACATACTTATCTCTGAACTCT
TTCTAACTAATGGTCAGTGAAGAAAAGTGCAAAATCCTTTAGTTTATTAGCTAATGCTTGGAAATGTAACTGTTCAT
TAATCCTTAATTAACTCAAGTAGCACTGAAGGAAAGGGTCAGAAACATTACTGAATAAAGTATAATAATCAATGACC
ACTTAATCCCAATAGCTCCCTAGAAGGGACAGATTTAGAAGGAAAGCGAAGACAATGAAATCAAGATGAATAAACAA
ATAACATTTCTTTGGAACTACTACCAAAAGTACATGACTATCTTCAGATTTGTTAAAGATAACATTGGGAAATAGAA
GAGTAATTTTTTTTATATATCTGATTTTAATATATTCTCAAAACCATTTATACACTACTGACACTGGTATTTCCGAG
CTATCAAAATAAACTGATAAATGATTCTTACTCAGTTTATTTCAAACTCACTGTTGCCACAAGGTGTCTTAGCAATT
TGATGAGATTACATTGCCTCCTTATACTACTAGATCATTTTAATTGCAACCTACCATTTAAATGACAATCCATGATA
TATCATCAGTCTTAAAGAGTCAAATCATTTGCTAGATTATAAAATAAACTACCTTATTTACTTTCTCTGCACTGCTA
CCTACTACAACGGAACAGCCACAGGTTTGCAAGTGTGAGCTGATGGCACTGTAAGTTAAAGAAAACAGATTAAAAAC
ATTGCCTATAAAACAATTTAACAAACTAAAAACAAAAAAAAGTAGGTGAGCTCTTCAAATAACTCAGAATAGCTTTA
TATGATAAACACCGAAGCTATAAGCACAATGTTATCTTTTATTTGTATAGGAACCTACATTTTCTAGAGACCTTTCA
CAGAAATTTTCTTATTGAGCCTTAAAACAGCCCAATTAGTCAGTATAATATCATTTAATTAATGTATTTATTTATTG
AAATACCATCATTTTATAGCTGAAGAAATTGACATGTAGAGAGATTAAGTGACTTACTTAAAGTCAAATGGGATTTA
AAATGATGTATGAAAGGCTGACACTGAACAGATACAGGACTAAAGTGCTTCTGATTCAAGCCATTAAGGCTCTTAGG
TTAAACACACTCATGCCTCTGATACTCCATCATGAGCCTAAAGGAAAAGACTGTGAACATAAAAGTGAATACTTTAT
ACTTTTACTTCTCTTTTATTAAAAGTAAAATTTCATGAAAATCTGTAACTGTGAAGAAACTTTAAAACAGAATATAA
GATAATACATGTAAAGCAACTAGTAAAGGAACTAACATGTAGGCACTCAACAAATACTGGCTATTTCTAGAAGAAAT

GTAAATAGGAAATGTTAGCTATGAGCTATTATTAAGTGTTTTTATGTTCCAGGCACTGTTCTAAGTGCTTTATATTA
TTTATCTTACTCAATGCTTATAACAACCCTACACATTAGGTACTATTACTATTATTGCCATTTTACAGATGAGGAAA
TAGGTGTATAGAGAATTCAGGCACCTTGCCCACGGGTACACAGCATTAATCCAGGGAGTCTGGTTTAAGGGCACAAA
CTCTTAAGTACTAAACTCCACTGCTGGATGGAAAAAGATCAGTATAAATATGAATAATTTTGTTCTACGCCTAAATA
ACTTAAGTTCATCTACAGTACAACTTAATATGAAAGGATTCTGTTAGCTTTAATGAGAAGTAAAACAAGAAACCAGA
ATCAAGCAAGGGGCCATGATTTCTTGTCTGGGATGGAAACTCGGTTTCTTTAAATAGCAAATGGAATAACACCAAAT
ATATATAGAAATATAATGAGTGAAAAATAACACAAATTTAAGCAACAGTTCAAATACGTAATGTCCCTAGAACAATC
TAAGTAGACAGTCTGTTATTTTCTTTCTTCCAAATCTTGTCATAGGTGAGCATAAGATGGTATCTGCTTCATCCAGC
TTTTATGAAAAGAAAAATTCTTACTTGAGAAGAAAGCCTTCATGACAGCTGTCACCAATATCATCATCATTGAGTAC
TGTATCAGCTATCTAAAATGCATCAAAAAATAAAAAAATTAGTCTGGCTGTAACATAGTGTTGAAATAACACTTTTA
ATATACAAGTTTTCGGAAGTCTGGATTCAATATAACACACTGCCTTCATTTCCGAGAATCAAGACTCCCCAAAAAAC
AATCTCTGTGCACTACCATAAACTTCAGAAGAACAAATGTGAAAGCTGGTCAAGCAGGTTAAACAATTTTTACAAGA
ACAACTTCCTCTTCTGAGCTGTCAGAATCAGGAGACTAACCTAAATGACAAAATCAGAAAACAACAAGAATAGTTTC
CTAAAGGTATCTCTTAACACTCATAGTGTGTGATTCAAAACGTCCTCAACAAATGATTAAGGAAACTAAATTTGTGA
CTACAAGTAAACTTCCATTAATGGTTACTACTTTGGCACACAGTTTTGTTTCAAAAGACACTACATTAAATATTAAT
TGCTCCTATAAGAGCTGGGATCCTCCCACTTTTAGGAATTATAAAAGTAATGAAATAAACAAAATGAATTTAATTTT
GTCATCACTGATCAAAAATGCCTCTGTTTTGCCATAAAATCCAGGATTTTGTGTGTGCTTATTTGCTAAAGTGGCTA
ATACTGTATGTGAATAGTATGTATGACAAAGTCCTTACTATTAAAATTAGAATATTAATAATATACATAATAATACT
ATAACCCCAAAAAACTCATAAAGTGTATAATTGCTCTCATTTAAACTTACATCTATTTCTTCAGGAACACTGTGTGA
TTTCATAGATGAAAGCAGTTCCATTACAGGAATCACTTCTCCAGTAAGCATTGGAATAATACTCTGACCCTGCACAA
TAAAGTGACATGAAGTGAAGAAAATCACGTAATATGAGAGAAGCTGGGCAATAAAAAATAAAAATAACATCAAACAA
TAACATTCTTTGATGAAAATACTTCGTAATTTGTTCAAACACAGTATCAAACAAGTCTACTACATGTCTAAAGGATT
TATATGCAATCCAAAGCTCACTTTTATTCTTTCTTTTTTTTTTTTGAGATGGAGTTTTGCTTTCATTGCTC
AGGCTGGAGTGCAATGGCGTGATCTCAGCTCACTGCAACCTCTCCTCCCAGGTTCAAGTGACTCTCCTGTCTCATCC
TCCCAAGTAGCTGGGACTATAGGTGCCGCCACCATGCCCGGCTGATTTTTGTATTTTTAGTAGAGACGGGGTTTCGC
CATGTTGGCCAGGTTGGTGTCAAATTTCTGACCATGCCCGGCCCTAAAGCTCACTTTTATTCTTTAGAGAGTATGGA
ATCATTGGTTTATCGTTTACTGTTACATGCAATGATTAAGTCATCATGCCTCTTTTAGAAAAGATCTCCTTTAAAAT
TTGAGATAAAAAAAATTTGTTAAAGGTCATCAATATATTTCATATTTAAAAATGAGGAAAACCAAGCACAAAAAGAC
TTTGAAATCCTTACCAAATAGGTAAGGAAAACTTGAATCAATACCTAACCTCCATACTCATAAAAGTATAATCTACC
CAAATGCAAATCAAAATCAGCACATATATTTTAAGAATCAATAAAACAGAAAAATTCCCTTTAGAGCTATTTCAAGA
TATTACTACTTATTACATCTTGAAATTGTAATTTTGAAATTTGTAGTCTATAGAATCAAACTGAAAATTCAGTATAA
CACATCACAAATGTAAAGTGTCTCAAATATGGATGGTCCCTCATTTATTCACTACCACCACCAGTCTCATCAGTTTT
[1139] GTGACCAACTTGGTTAAGTAAATTTTTTGAGATATAAATGAAATTGTCAAATGACTATGCATTTTTAGCTAAACATA
TTTTTTAAATCATACATTATTTAAGATGAATTTAATACTAGTTGTTTTTCCTCACTTATTTTATGAAATGATTTTAC
TCAAAGTCTTCATAAGCATCTTTAAGTTAGAATCTTTGTCAGACCCAGGGCCATTTTTGGAGTAACCTTAACCAGTT
TTTCAGAGCCCCATATTATTAAGTTGCTTGAGAATTTAAATGTGATGCTACTTCTGGAAGTTTTATCCTAAGCCATA
TGCCCATTTGCATAATGCTGAAAGTTTTATTTAAAAAAAAACCATCCTTTAGTAACCTCCACAACTAACTATTCACT
GTTTTTAGTTTTTAAAGTAATAATTATCATGCCTGTTTACAATTACAATTCACACATTCAATCTAACAAGAATAATG
ACTAGATCCGTGTTAAATTTCCTTCCCTGTGAAGCAATTTTATCAGATGACAGCTACAACTGAAGTTGTTTCAAACT
AATGCATCATCCCCAAACAGTATTGTTCAAAATAAAGTCGTTGTGAGATTTGCAAGAACTCAATCAAAAGGCAACTC
CTCCTTTTCGGGAAGAATAATTTTGGGAAAATATTTCCTCTTAGGTTTAAGCATACATAGTATTTCATTCACAGTAT
CTCAGACATTATCAGTATAAGTGAATGAATAGCCTCACTGAAGCTCAACAACACCAAAAAAAAAAAAAAAAAAAAA
ATCCTACAGGGCTAAATACAGAAGAGGCTCTAAAAGAAAATCTCTTAAGTTTCTATTCCTCCTTGTACTTCCCAAAC
TTGAACTTCTCAGCAGTAAGATAACATTTTTAAGAAGAGCACTTAAAAGAGAGACCAAAATTCATTAATAGTAGTCA
ACTTAAGTAAAGGTTTCTGGTTTGAAAAAACAAAATCCCAGTAAAAGCAGAATTTTAGTTGGTTCTAAGTTTCCTCA
ACTTGCGATAAGTTTACTTAATTAGTCTACTAATAACTAGTGGGTTAGAGGGTGCTGAAAGTTACCCCATTCCTGGG
GACCCTGCTTATTGACCAGCAAATAAGGACTGGGATTCTTTGGGTAAAGGGAAATCTTTTCTTGTTAAGTCAGACCT
TTACACAGAATAACTGTCTCTGAATTGGAAAGCTATCTACAAAAGTACAAACATAACAATTTGGTAAAGGAGATCAT
TGTATTGGGTTCTGTATTATGGCCATGTATTTTCACAAGTTTTTTTTTTTAATTACTTTTTTAAAGTATCATCTGTC
TCATTCATGCTAAAAAGAAGCAAAGAAAGGCAAAACAGCCATGTTTAAAATATTGGAGTTTTACAAGGAGCATTGAG
GGTCACCCACAAGAGGAAATGGAAGTAAAAGTGAAGAACTCTTTCTTCACTGGAGATTCTCCTTCAAAAGAACTTCT
CTGCTTTACAGTGAAATAGTCTGTACTTAGTTTCCGCAGGGGAAGCCACACCCTTGTAACCATGCTTCTCAAACTCT
TAGTGTCTGTTCCTGAGGGGCATTCAAAGCCAAGGGATAAACATGGCACATTTTCCTAGAGGAGAGGGTAAGAAATA
TCACTGACAAATTTTAATACTAAAATAGTTATGGAATAAAATGTAAATTGCATGAGTCTTAACGATACAACATAAGA
CTTAGAAGAAATATTGTGTGGACCTGGGCCTACACCCCAGACAGATACCTCAGGGGTACATATGCTCTCCTTCTGTT
ACAGCTACTTCTAGGGAAAGGTTCGAGAAGTAGTACCTTAAAGAACATATCAGAGACAATTTTTTTTATTTTTACTA
TGAACAAGTTATCCAAAATTTATTCTGGGCAAACAGAAAAAAAAAGGGAGCAAATATTAATTTGTAGATGCAATTAC
TATTTTCCTTTGTTTACTGATTTAACTCTTTGGGTTTAAGATATGGAAATCTTCCTCCAGTTTATTCTGTACACCTC
CATAAAAGCTCCATTAAAGGCTTATTCGTATGTCTCCAAGGCCTTGACAAATGTAGCCATCAACCTTATACAGATAC
ATGCTGTGAGAAAAACATTTGACAGTATGCAATTTGCATATACCTGATCTTCCATTCTCTCTGTGCCTTCTAAGATA
ATCTTCTGGACATTTTCTTGTCTTTCCTGAGCAAGAGAAAATTTATTTAAAAAAACAACCCACAACATTTTGATACT
TGCTTATTTTTCAATAGACATGTTCTTGTGTAGTAATTTAGTTCACAAGAAAAATACTTTCTACTTTAGGGAAAAAA
TGGGGGCAGGGGTAGGAAATTAACCCAACAAATGCATGTTCTCATAAACAATACAAAATAAAATCAAAACAACCTTT
ATTCTGCAGTGAAAAAAAGATAACTTCACAGAAAACAGTCAATGTAACATCTGCATAGTTTCAAAAAGGAAAAGAAT
GACTTGCACTTTTCAAATTAAACATTATGATGTTGTTTAAAAGATTCTCCTGATTTTAAGAGTTTCATAATGTGAGA
AAAAAGGAAGTAAGCCTGCAAACATAGTAAAAAATTATTCTTTTAAAAGATATTATTTTTCCTTACTATTGGGCAAA

AGCCTTTTAAAACTGGTAATGCTTAATGGACTTTCAGGTTAGTATCAAACTGGAACACAGGAAGGAGAATTCAATGT
GTTCTTTAGATACATCAAAACTATACTGAAATGTAAATAGCATTATATATTCAACTACAGGATTTAGGAAAACAATA
ATTTCTGTAAGATTAAAAGGAATTCTCTTGGGAACCATTCCATTCAACCTCCTCATTTTATGAACCTGGGAACTTGG
CAAAGAGGTTAAAGAGACCAAAGGCTACATGACCAACAGCTTATGAAACTATTACTTTGAACTGTTATACTTACACA
TAGTAGTAAGCAAAAGACAGAATTGTGCAATGAAAGGGAAACAAAAGGTATTAGAGTCAAAGGCTCCCAAGAAGAAT
CCAGGGTCTAAAAGTTTCTTTATTTGTCTAAGCTTTAGCTTTTCATCTATAATGTGGAGCTACCATTTCGTACCTTC
CACAGTTAATATGAAGATGACAGGTATCAGACCAGATGTATTTGTATCTAATAGGGTAAATGCAAAATAAATAACAT
TTATTGTTTGATGTTCACTGCATATAATTAAAAAAATAAGATTTATATGTACCAGAAAATAAGCTTTCAACAGATAG
GTTAACATGATTAATAAGCTGAAAAATCACTTACCTTATGCATCCATATTCTTCCTTTCCGGATTATATGTGTTAAT
CTATCAACACACACTCTATGAAGTGGGAGGTAGAAACTAAGTTCTGTCTGTGGAAGTATAATTGATAGTCCATATGT
GCTGCGATCCCCATTCCAGTTTCCATCAAAGATTAATGAAACAATAATCACTCCCTTTTCAGACAAGACAAAAAACT
TTACATCTATAGCACCACTCTCTGCATTTCGAAGGATTTCTCCATTTAGAGTGTGGTTGGCAAGAAAAGTTATTTCT
CCATCACTGAGAAGTACCTGTTCTGTCTTTGGAGCCCAAATGTGCCTTACTCTAGGACCAAGAATATTGTCCCAGTA
AGCAAAAGTAGCTGCTAATAAAGGTGATTTGCCACTTAAAGCAATCTCTGTCTTGGCAACAGCTGGAGATGGCGGTG
GGCAAAGAGTCGACATCACTGCATTCCAACTGTCACATTATCCAAATGCTCCGGAGATATCTAAACAATGACATATG
AAACCAATGATTAGGTTCAGCAATTTAAAGATATCCATCAAAACCCCAAATGATTTAGACATATTTGGTTTGTCCTC
TTAAGTCAAAGATGTGGAATCCTGTTATCTCCTATCAGGATAAAGACATTCAACTAGCACAGTAGGTGCACATTAAA
TGTTTGTTGATATGATCATTTTACAAGACATGGTAACTTGTTACTTATATTCAGGGCATACATTTAGAAATTCAAAG
AAATAACTTAAAAAAGGGCTTCTTTACACTGATATTAAATGTTACATACTAAAGCTCATAGAATAGACCCGCAGTAT
TCCCAAATATCCAGTCCATGTGCAATTCTAGTATGACTGGAGATTTGGCCCCTAACCCATAGCAACTAAAAAGGAGA
AAAACAGGAAGGGAAAGGCTCAGCTAGAGACTGACACTTGTGGGTTGAATTGTGTCCCCCAAAAAGATATGTTCAAT
TCCTAACCCTTGGTGTACGTGAATGTGACCTTATCTAGAAATAAGTGTAATCATGTTAAAATGGGGTCATACTGGAT
TAGAGTGGGGCCTAATCCAATAACTGCTGTGTTTATAAGGAGAGAGATTTGGAGACACAGAGACAAATGGTAGACAG
CCATGTGAAGACAAAAGGCAGATACTGGATTGTTGAAACTACAAGGCAAGAAAGGAACACTCAGGATTGCTGGTAAC
CACCAGAAGCCAGGAAGAGGCAAGGAAAGAGTCTTCTCTCTTGAAGATCATGCCCCTGTCAACACTTTGATTTCGGA
CTTCTAGCTTCTAGAATTGTGAGAGAATAAATTGCTGTTGTTTAAAGCCACTCAGTTTGTGGTGCTTTGTTAAGTAA
TCTTAGAAAAGTAATACAACACCTAACAACAGAAATACTTTAAAGCCGCTAAAAGGTCAAAAAAAAAAAAAAAAAAA
AAGACATGGAAATACCACAAGTCTGGAGCCATAACAAAAAATGGGCAAACAGTCCTGTATCCTCAGTGAACTCTCTG
GTTATGAGAATACTGAAGCCCGATCCTGATGTTTAAAACGACATTGAAGTATCAAGACAAAGATAAAAATATTTAAT
ATGCTAGCCAAGAAACCAATACAGCATTTCATCACTGCAAAGAGAGTTCTACACTAAATGGCTAGAATTTAAAAGCT
TTAGTTATTTAGAACACGTAGAAAACAGAAGGGCTAAATAGGGCCCGTTCAAGCCTTTGAATTTAATGAGAAAACAG
ACATGAGGAGAAGAACATAAACGCTCACATCCAAGACAGAACCCAGGGCTCTTGGTCCCCTTGCTCAACTTGTACAT
[1140] CTTAATCCACATAAACATACCACTCTAAAAAGGTACATCCTATGTGATATTAATGTAAAACAAATCATTCTTGCAAA
TACAGTTATGTGCCATGTAACGTTTCAGTCAATGGTAGACTGCATATATGATGGTAGTCCCATTAGATTACAATGGA
CCTGAAAATATGCTATTGCCTTAGTGACACTGTAACCATCATAAGGTCTTAGTACTATTTTGCAAGTTATTTAAAGT
ATAGCACATACAATTATTACAGTGTACAACACTTGATAATAAACTACTACATTGCTGGTTTATGTATTCACTATACT
ATGCCTTTTATTGTTATTTTAGAGTGCACTCCTTCTACTTTTTTTTTTTTTAAGTTAAATGTAAAACAGCCTCAGGC
AAGTCCTTCAGGAGGTATTCAACAGAAAGCACTGTTATCATAGGTGACAGCTACATGTGTGTTATTGCCCCTAAAAA
CCTTCCAGTGGGACAAGATGTGGAGGTGGAAGGCAGTGAGGTGGAAGGGAGTGATACTGATGATCCTAATCCTGTCT
ATGCCTAGGTGAAAGTGTGTGTGTTTTAGTTTTTAACAAAAACGACTAACAAGTAAAAAAAAAAATTTAAAATAGAA
AATAGAAAAAAGCTTCTAGAATAAGGATACAAAGAAAAAATATTTTTGTATAGCTATACAATGTATTTGTGTTTCAA
GCTAAGTATTTTAAAAGTTAAAAAATTAAAAAGTTTACAAAGTTAAAAAGTTATAATTTTTTATTGAAGAAAAACTG
TTAAGATAAATTTGGTGTAGCTTCAGCGTACTGTGTTTATAGTCTACAGTGGTGTACAGTGTTCTAGGCCTTCACAT
TAATTCACCACTCACTCACTGACTCACCCAGAGCAACTTCTAGTCCTGCAAGCTCCATTCGTGGTAAGTGGCCTACA
CAGGTATACCATTATCTTTTATACCATACTTTTACTGTACCTTTTCTCTGTTTGCATATATTTAGATAAATATTTAC
CACTGTGTTACAACTGTCTATAGTATTCAGTACAGTAACAGTTGTACAGGTTTGTGGCCTAGGAGCAACAGACTATA
CCATACGGCCTAGGTACATAAAGGCTATACTATCTAGGTTTGTGTAAGTACACTCTATGATGTTTGCATAATGACAA
AATCGCCTAATGATGCATTCCTAAGCAATGTGTGATTGTACTATAATTGAAGACTTGTTATCTAAGACTGAAAGTAA
AAAGAATTGCAATTTCACCTAAGCAAGTCTAAAACTGTGAAGTCTATTTATAATAATAGCAATACAAAGCAGCTAAT
AGGCAAACTATGATATACCTATCTTTGCCATATGATTGCTTTGGGAGCTAACATTTGATCTGTAAATGTATGACAAA
GTAAACAATTTTACTTAAAGAATTTCATCCACATCTTGTCAAGAGAGTTCAGTCTGATGGAAAGCACTGACTTCTAT
TTACAGAGCATTAGATGAGTGCTTTTATCATATTATGAGTAGGCATACAGAGCCTGGCAAAACAGTTAACTCTAAGT
ATGTACAGAAATGGTTGAACACAACGACAGTTTTAACACGTGTATTTGTAATTTCAAAAATTCATTTAGGTAATATT
TACTTTTAAATATGTTGTATCAATTTAATAGTCTTAAGAGACAGCACTAGATATAAGCCGTACAGCTTCTTTAAAAT
ATCCACTGTTTTTAATACAATGTAAGCAGTCAGTTTACAATGATCAAATATAGGAATGTAATCTGAATTGAAATGGT
AATGACACTACTGCTGTCATAACTAACAACAGCAAACTGGAGGCCAACATAATGAATTAAGTTAACATACAACCATA
AAATTATATTGCAAACATATTTTTCTTTCATTCTTTTAGGTTAAAAAGGTGGATAATCATAAAGGCAATATTACAAC
TCTAATATTTCATCATTAAACTGAAAATAAAAGTATTTCCTAAAACAGAACTGAACCCTGGAGCAAAATCTGATTGA
ATTATAGGGAAACTTTTACCACGTTGTGAAAATTGAACTATTATACTGCTAGTTACACTCTCACTCCTAACAGAATA
AGAAAAAAAAAAATGGGCCGGGCATGGTGGGTCACACCTGTTATCCCAGCTCTTTGGTAGGCCGAGGCAGGTGGATCA
CCTGAGGTCAGGAGCTCAAGACCAGCCTGGCCAACATGGTGAAACCCCACCTCTACTAAAAATACAAAAAATTAGCC
GGGTGTGGTGGTGGACACCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCTCTTGAACCCGGAGGTGGC
AGAGGTTGCAATGAGCTGAGATGGCGCCACTGCACTCCAGCCTGGGCGACAGAGAGAGACTCTGCCTCAAGAAAAAA
ACAAACAAACAAACAAACAAAAAGAATAAGAAAGAAAATGAAGGACAAAGATCATACTGAATTGCTTAGTTTTAAAT
CCTACCAAAAGAAATAGCCTGGGAAATGAAATGTCACAGAGAAGTATAATCAGGAGAGCTGTACAATTATTTTACTA
ATACTTGAAGTCATCGTCTTTGGTGAGAAAAATCCATACATGCAAATGCAGCTGAAAAAAATCAGCTCAAAACCAAT

[1141]

AGTTGTTTATGTACCTATCTTACGTACATGTAGTGCTGTCTACTCCAGAGAGTTACCAAACATTAGCCAGTCTTTTG
AGGGAAGCCAAGATTCAAATTGAGTGAGACGGTGGCTTGCTCACAGGGTTCATGAGAGGTTTCCCAATACACTTTCT
GGAAATAATCCCATACATGCAGACATGATTACATTAATTAACATCTGCTAAAACTGTTAGTAGAGTGCTAAGTTTGA
GGTTTTGCTTTTTCTTTAAACGTCTGTTAAAAAATCAACCATCTCTTCCCTGATTGGTATTTAGAAAGGTGGTTGGT
CCACTGCTATTGTAGTGAAAATTCTACAATCATAAAGCCCTCACTTCTTGTTTTTTAGAGACAGGGTCTCGTTTTGT
CATCCAGGCTGGAATGCACTGGCAGGATCATAGCTCTCGGTAACTTCAAACTCTTGGGCTCAAATGACCCTCCTGCC
TCAGCCTCCCAAGTAGCTAGGACTACAGGTGCACATCACCACGCCCGGCTAAGTTTTTAATTTTTTGTAGAGACAGG
GTCTACGTTGCCCAGGTTGAGCTTGAACTCCTGGCTTCAAGTGATCCTCTTGCCTCCGCCTCCCAAAGCTCTGGCAT
TACAGGTGTAAGCCACCTTCTCCAACCTGGCTCTCAATACTTGTAACCATGCTGTTTATTTTCTCCCAGCCCAAAGA
GAAGCAGGATCCTAAACCGTCCACTTTCCACAACAGGAGCTGCCCAGGACCACTTCAAGGACAGTGAACTGTTTACA
GTACCAGAAAGTTCACAACACTTTCTCAATCTTCAACATCAGGGAAGACTGGAAGGTGAAGTTCATATCACTATCTG
GCCATTTCTCACAGTTCCAAGTTTCTCAGACAATAGGTAGGCTAACCTAGTCCTCCTGGGAACTATCTAATTAACGT
AGAATAGAACCCGAGGGCAGACTTGAAAAACAGAAGTCCTCCTTGGTTTACTTTGTTTCTCTGAAAGCAAATTGTGG
AGTGCCAACATAGCCAAACAAAATATTTTATCAACTTCATAAGGTGCTTGTAATTTTTTCCTGGAGCAGGTAAATGC
TGGCTTAGTGAACAATCTGGAATGTGGTAATTACTCTCGTTCTTGTTTCAGATGTACTATCAGCATGTAGCAGTTTC
CAACTGATTCAGGGTTTTCCTAAAGTGGCAGGCCTTGGCAGAGGTGGTGACAACAATGCCCGTGTCAAATGACACCG
TATTTCAAGTATTCTGACTCCAGGTTATTAATATCCCCTATATGATAGTCTTGTTTCTGTGATATTCACAGATTATG
TTAAAAGTTTCCCAAAGTCTGAGAAAAATCATATCTTAACAGTATCTTTTTTTTTTTTGATCCTTTGTACAAAAGTA
GAAGTAATGCCAGACAGATTACGTACCCTTGTTGTGAACAACTGGTGCATGGCAACTGTTTGAATAGAAATTTACCA
ACTGCCACAACCAGGCAACTACTCTCCCAGAGCCTAACAATCTCGATTGCATCTGAAAGGGCCACCCCTCCTGGGAA
AGTGCAGGACCTCCCTCCTGTTTCTGAATACAAAGCCTGGTGGTGTTCAACGCGGCCAGATAGACCCAATGAGCACA
CGGACATGTAATCTGTGCACTTCTTTAGACAACTGATTACCATCAGTCAAGTGATGCCCAAGTCACAATAGTCACTT
CCTTTAAGCAAGTCTGTGTCATCTCGGAGCTGTGAAGCAACCAGGTCATGTCCCACAGAATGGGGAGCACACCGACT
TGCATTGCTGCCCTCATATGCAAGTCATCACCACTCTCTAGAAGCTTGGGCTGAAATTGTGCAGGCGTCTCCACACC
CCCATCTCATCCCGCATGATCTCCTCGCCGGCAGGGACCGTCTCGGGTTCCTAGCGAACCCCGACTTGGTCCGCAGA
AGCCGCGCGCCGCCCACCCTCCGGCCTTCCCCCAGGCGAGGCCTCTCAGTACCCGAGGCTCCCTTTTCTCGAGCCCG
CAGCGGCAGCGCTCCCAGCGGGTCCCCGGGAAGGAGACAGCTCGGGTACTGAGGGCGGGAAAGCAAGGAAGAGGCCA
GATCCCCATCCCTTGTCCCTGCGCCGCCGCCGCCGCCGCCGCCGCCGGGAAGCCCGGGGCCCGGATGCAGGCAATTC
CACCAGTCGCTAGAGGCGAAAGCCCGACACCCAGCTTCGGTCAGAGAAATGAGAGGGAAAGTAAAAATGCGTCGAGC
TCTGAGGAGAGCCCCCGCTTCTACCCGCGCCTCTTCCCGGCAGCCGAACCCCAAACAGCCACCCGCCAGGATGCCGC
CTCCTCACTCACCCACTCGCCACCGCCTGCGCCTCCGCCGCCGCGGGCGCAGGCACCGCAACCGCAGCCCCGCCCCG
GGCCCGCCCCCGGGCCCGCCCCGACCACGCCCCGGCCCCGGCCCCGGCCCCTAGCGCGCGACTCCTGAGTTCCAGAG
CTTGCTACAGGCTGCGGTTGTTTCCCTCCTTGTTTTCTTCTGGTTAATCTTTATCAGGTCTTTTCTTGTTCACCCTC
AGCGAGTACTGTGAGAGCAAGTAGTGGGGAGAGAGGGTGGGAAAAACAAAAACACACACCTCCTAAACCCACACCTG
CTCTTGCTAGACCCCGCCCCCAAAAGAGAAGCAACCGGGCAGCAGGGACGGCTGACACACCAAGCGTCATCTTTTAC
GTGGGCGGAACTTGTCGCTGTTTGACGCACCTCTCTTTCCTAGCGGGACACCGTAGGTTACGT

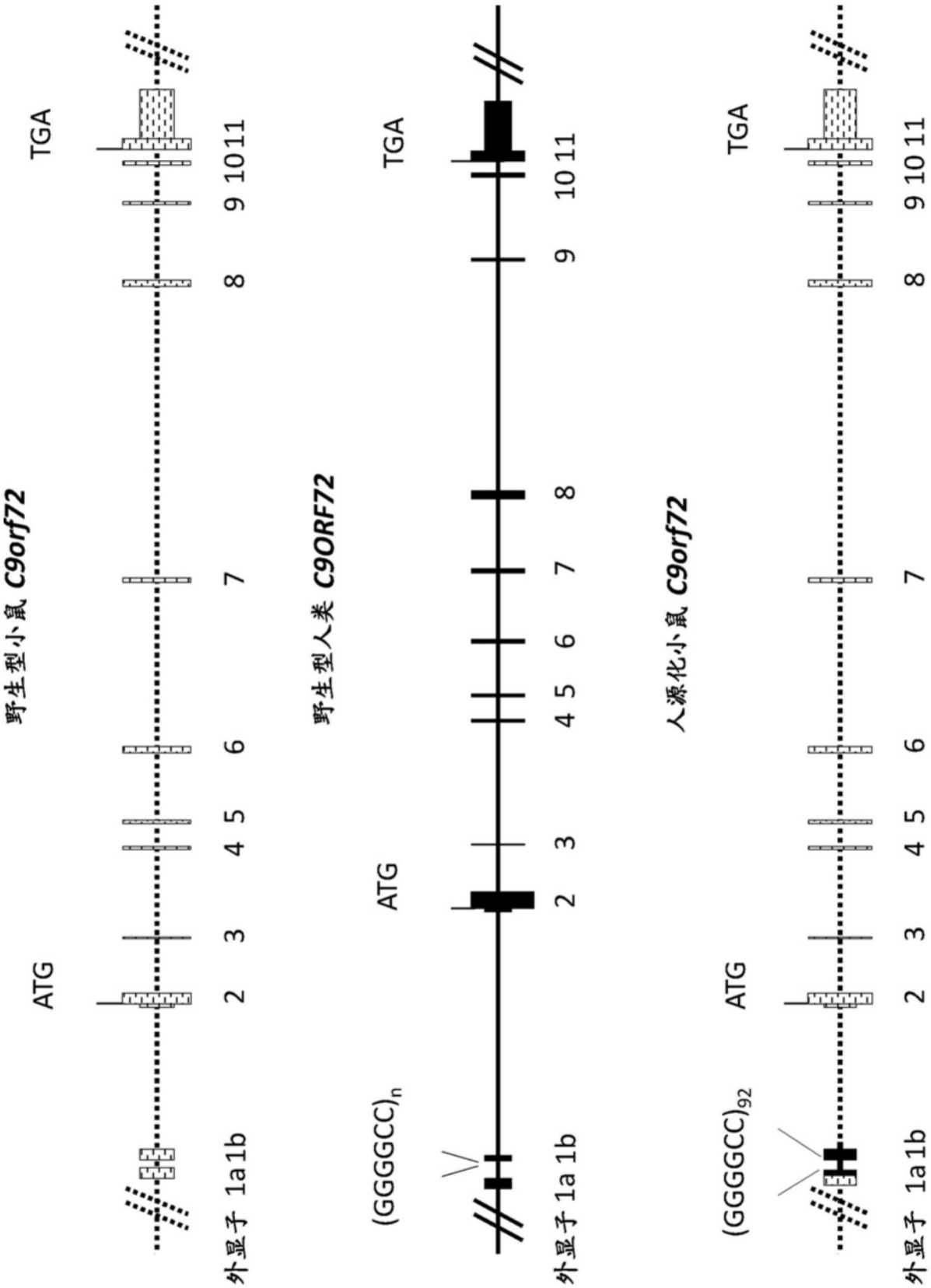

野生型小鼠 C9orf72
ATG
TGA
外显子 1a 1b 2 3 4 5 6 7 8 9 10 11
野生型人类 C9ORF72
(GGGGCC)n
ATG
TGA
外显子 1a 1b 2 3 4 5 6 7 8 9 10 11
人源化小鼠 C9orf72
(GGGGCC)92
ATG
TGA
外显子 1a 1b 2 3 4 5 6 7 8 9 10 11

图1

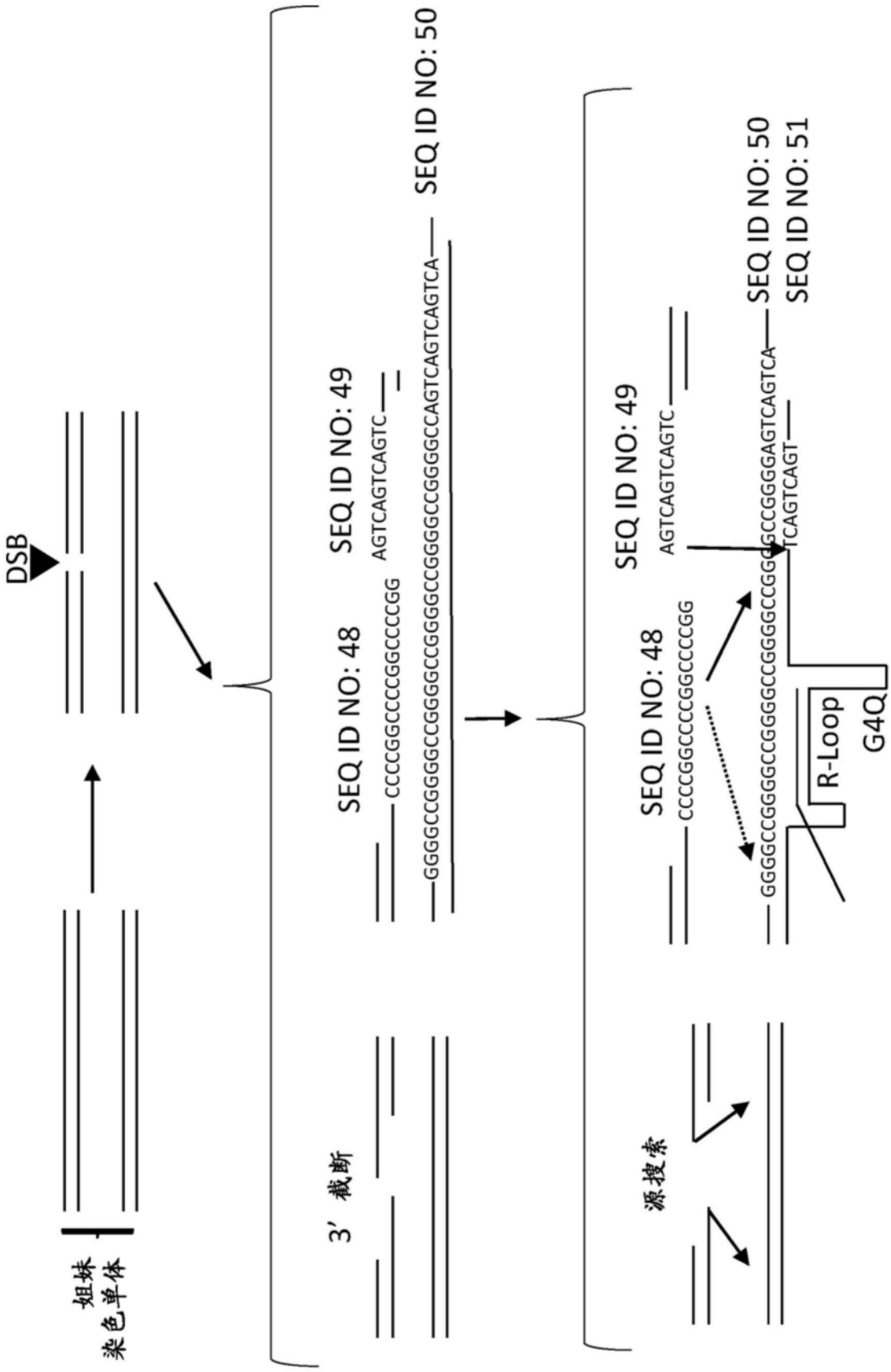

DSB
姐妹
染色单体
3′截断
SEQ ID NO: 48
CCCCGGCCCCGGCCCCGG
SEQ ID NO: 49
AGTCAGTCAGTC
SEQ ID NO: 50
GGGGCCGGGGCCGGGGCCGGGGCCGGGGCCAGTCAGTCAGTCA
源搜索
SEQ ID NO: 48
CCCCGGCCCCGGCCCCGG
SEQ ID NO: 49
AGTCAGTCAGTC
SEQ ID NO: 50
GGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGAGTCAGTCA
SEQ ID NO: 51
TCAGTCAGT
R-Loop
G4Q

图2

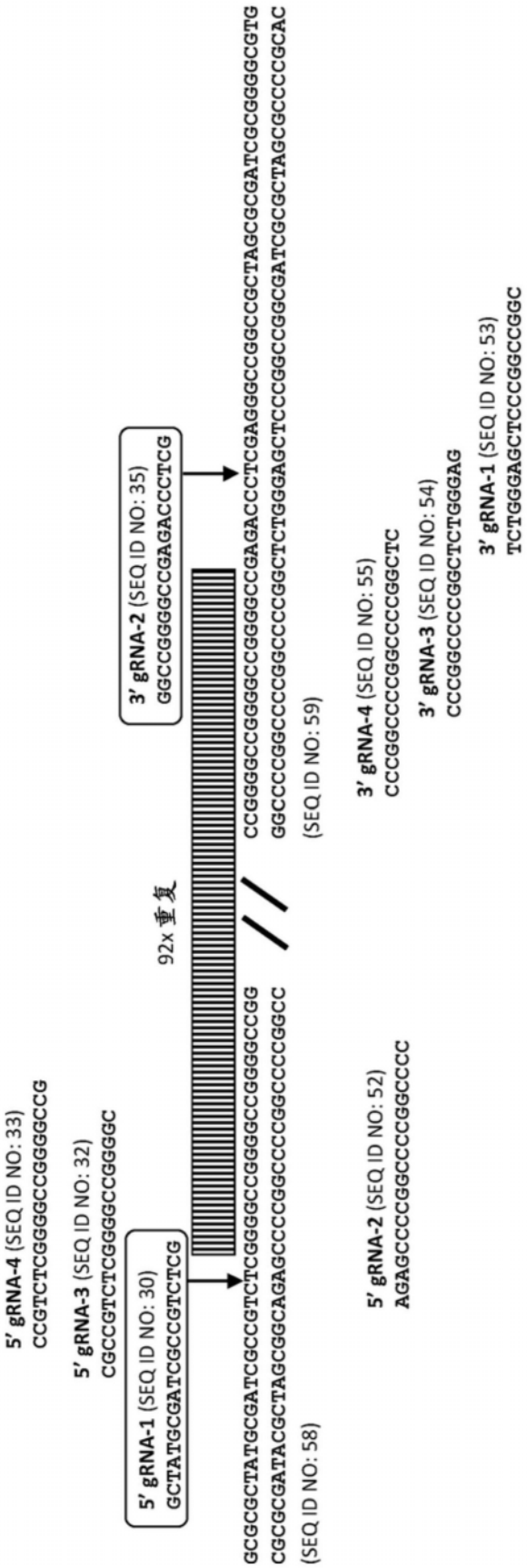

5' gRNA-4 (SEQ ID NO: 33)
CCGTCTCGGGGCCGGGGCCG
5' gRNA-3 (SEQ ID NO: 32)
CGCCGTCTCGGGGCCGGGGC
5' gRNA-1 (SEQ ID NO: 30)
GCTATGCGATCGCCGTCTCG
92x 重复
3' gRNA-2 (SEQ ID NO: 35)
GGCCGGGGCCGAGACCCTCG
GCGCGCTATGCGATCGCCGTCTCGGGGCCGGGGCCGG
CGCGCGATACGCTAGCGGCAGAGCCCCGGCCCCGGCCC
(SEQ ID NO: 58)
CCGGGGCCGGGGCCGGGGCCGAGACCCTCGAGGGCCGGCCGCTAGCGCGATCGCGGGGCGTG
GGCCCCGGCCCCGGCCCCGGCTCTGGGAGCTCCCGGCCGGCGATCGCGCTAGCGCCCCGCAC
(SEQ ID NO: 59)
5' gRNA-2 (SEQ ID NO: 52)
AGAGCCCCGGCCCCGGCCCC
3' gRNA-4 (SEQ ID NO: 55)
CCCGGCCCCGGCCCCGGCTC
3' gRNA-3 (SEQ ID NO: 54)
CCCGGCCCCGGCTCTGGGAG
3' gRNA-1 (SEQ ID NO: 53)
TCTGGGAGCTCCCGGCCGGC

图3

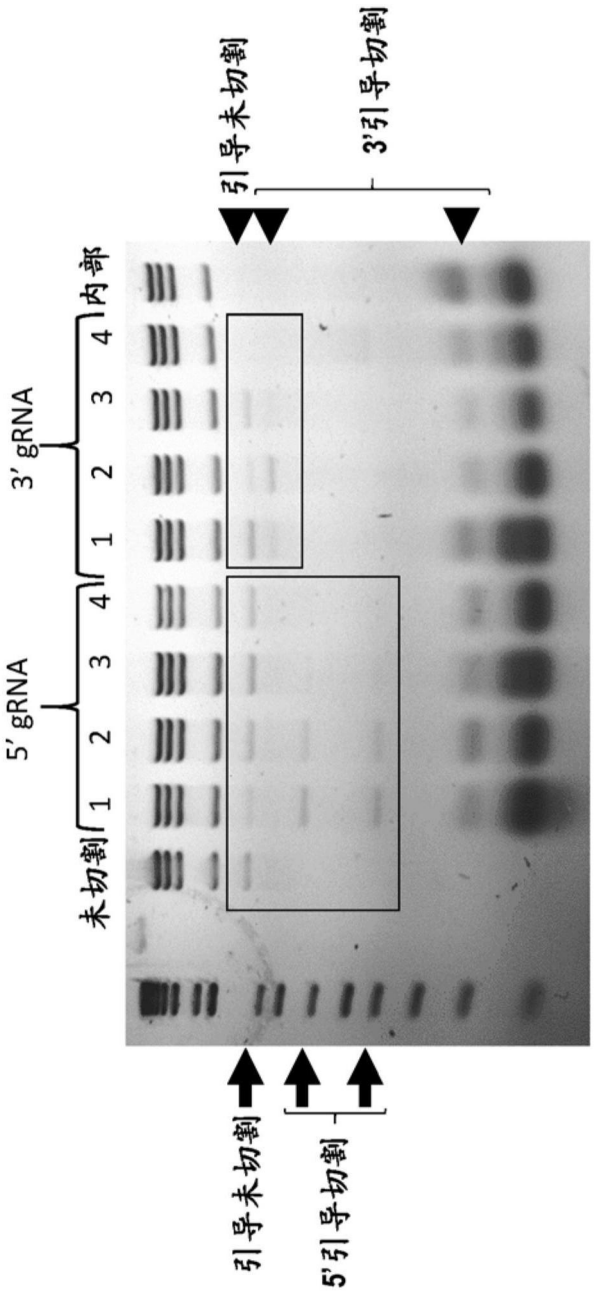


图4

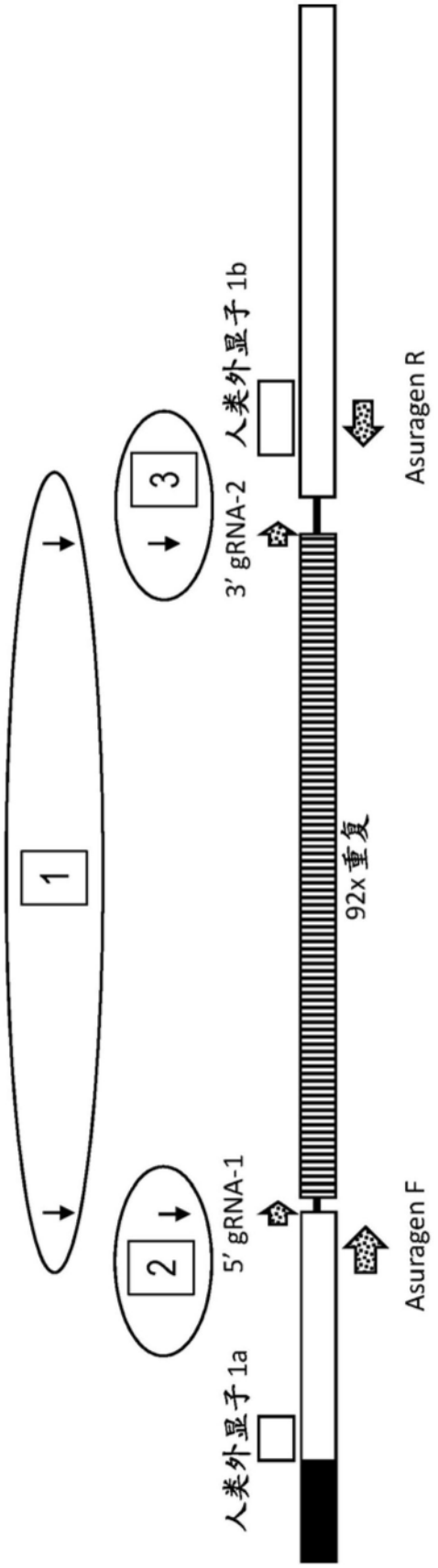

1
2
3
人类外显子 1a
5′ gRNA-1
92x 重复
3′ gRNA-2
人类外显子 1b
Asuragen F
Asuragen R

图5


基因特异性 PCR (2-引物)
GGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCC (SEQ ID NO: 60)
CCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGGG

图6A

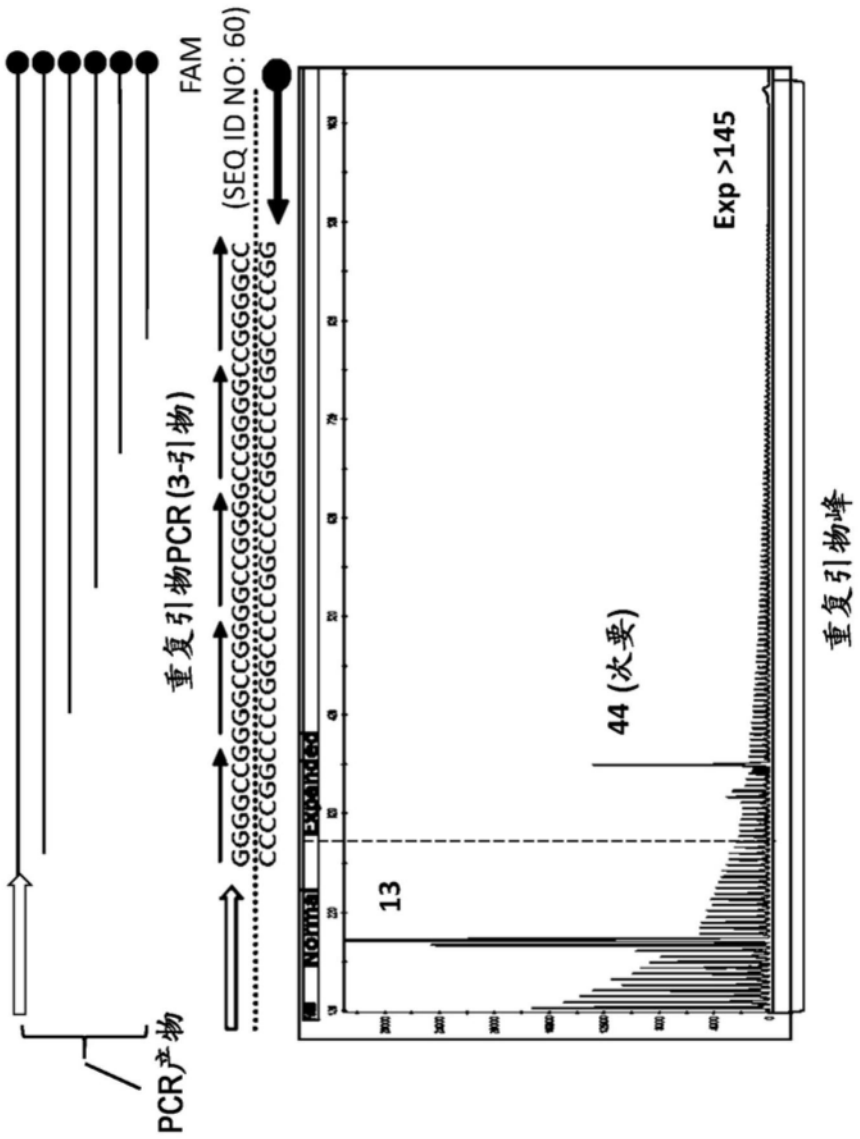

FAM
(SEQ ID NO: 60)
GGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCC
CCCCGGCCCCGGCCCCGGCCCCGGCCCCGGCCCCGG
重复引物PCR(3-引物)
PCR产物
Normal
Expanded
13
44 (次要)
Exp >145
重复引物峰

图6B

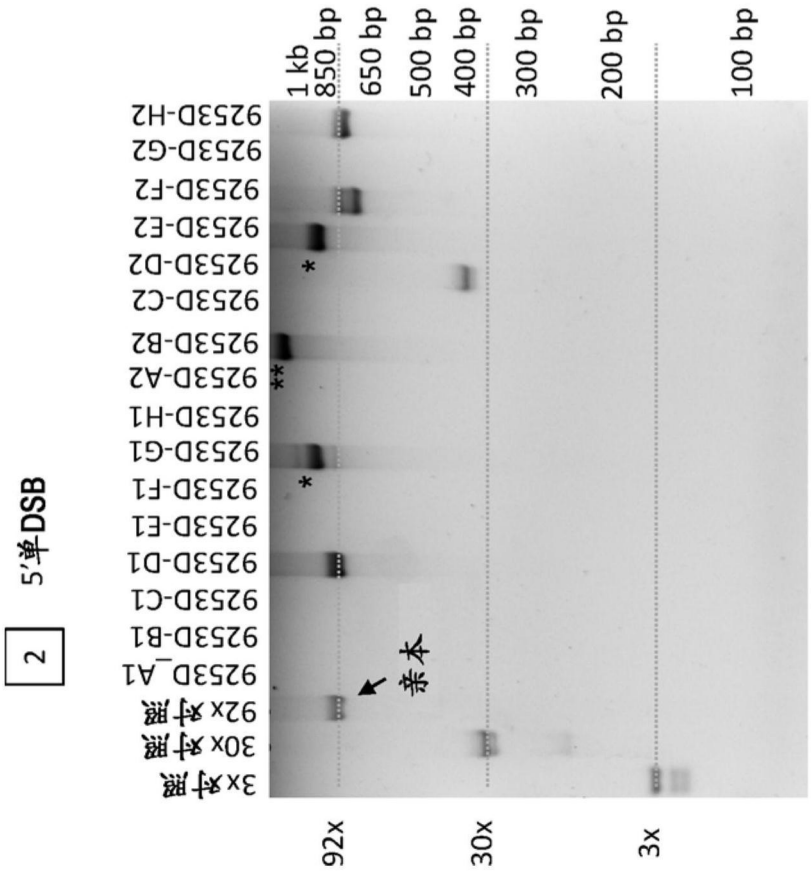

2
5′单DSB
1 kb
850 bp
650 bp
500 bp
400 bp
300 bp
200 bp
100 bp
9253D-H2
9253D-G2
9253D-F2
9253D-E2
9253D-D2
9253D-C2
9253D-B2
9253D-A2
9253D-H1
9253D-G1
9253D-F1
9253D-E1
9253D-D1
9253D-C1
9253D-B1
9253D_A1
92x 对照
30x 对照
3x 对照
亲本
92x
30x
3x

图7A

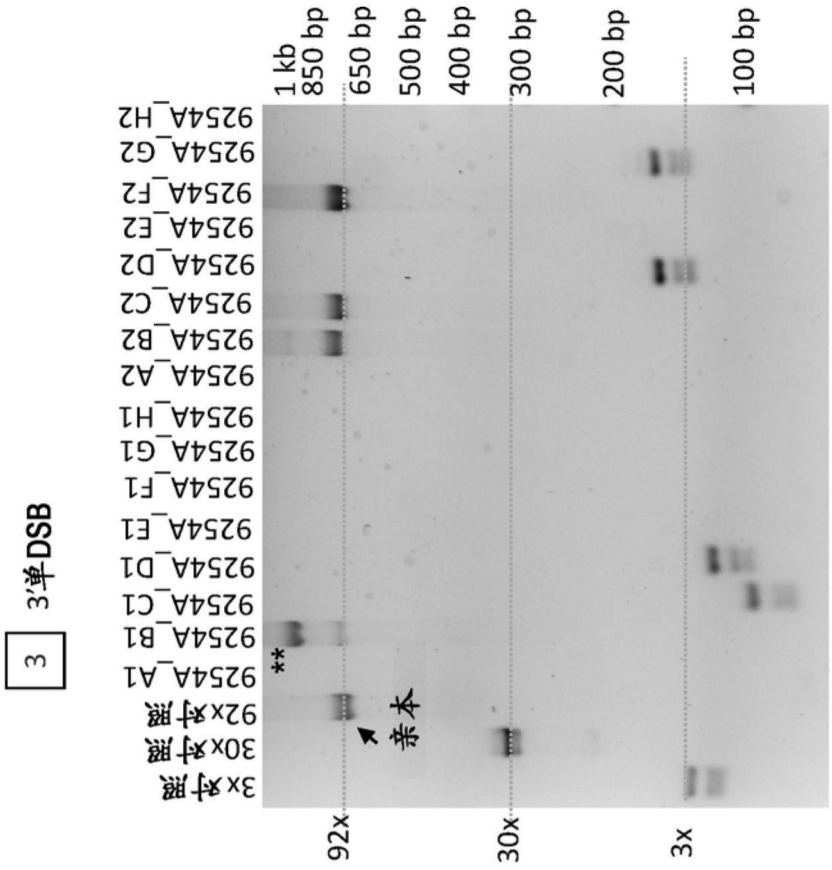

1 kb
850 bp
650 bp
500 bp
400 bp
300 bp
200 bp
100 bp
9254A_H2
9254A_G2
9254A_F2
9254A_E2
9254A_D2
9254A_C2
9254A_B2
9254A_A2
9254A_H1
9254A_G1
9254A_F1
9254A_E1
9254A_D1
9254A_C1
9254A_B1
9254A_A1
92x对照
30x对照
3x对照
亲本
**
3
3'单DSB
92x
30x
3x

图7B

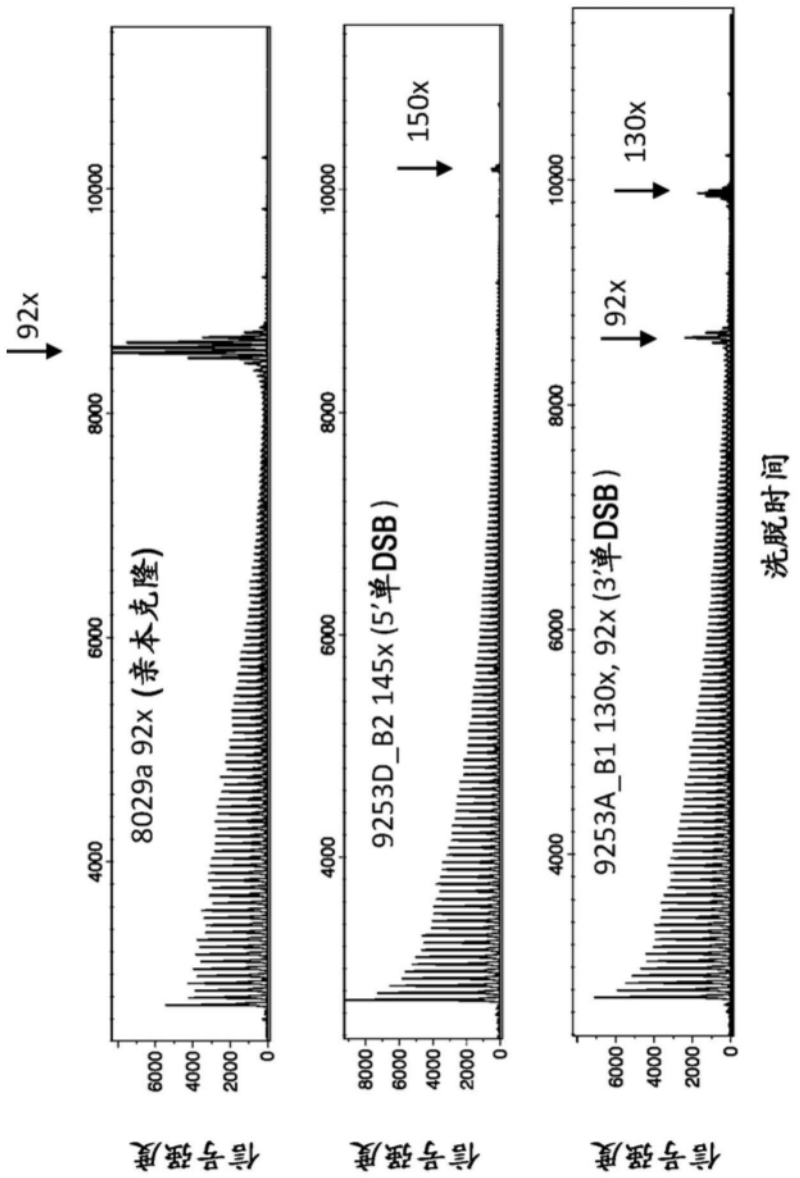

92x
8029a 92x (亲本克隆)
9253D_B2 145x (5'单DSB)
150x
9253A_B1 130x, 92x (3'单DSB)
92x
130x
信号强度
洗脱时间

图8

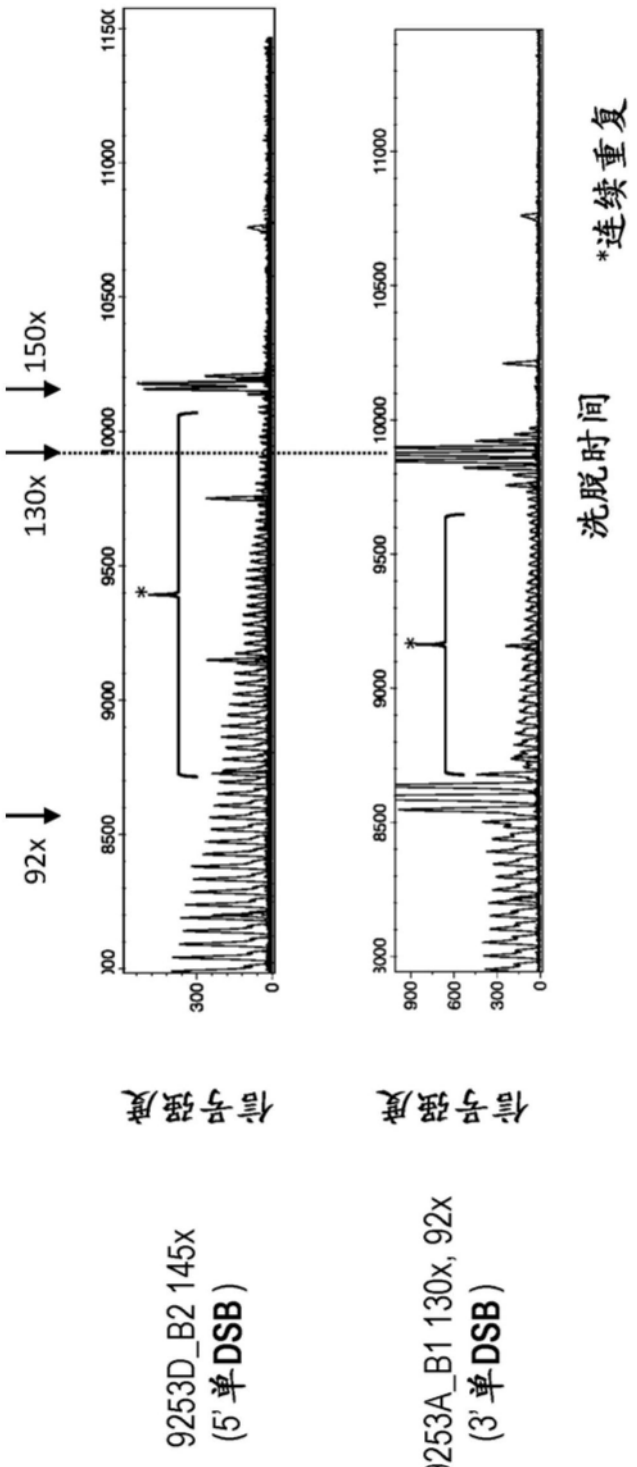

92x
130x
150x
信号强度
洗脱时间
*连续重复
9253D_B2 145x
(5'单DSB)
9253A_B1 130x, 92x
(3'单DSB)

图9

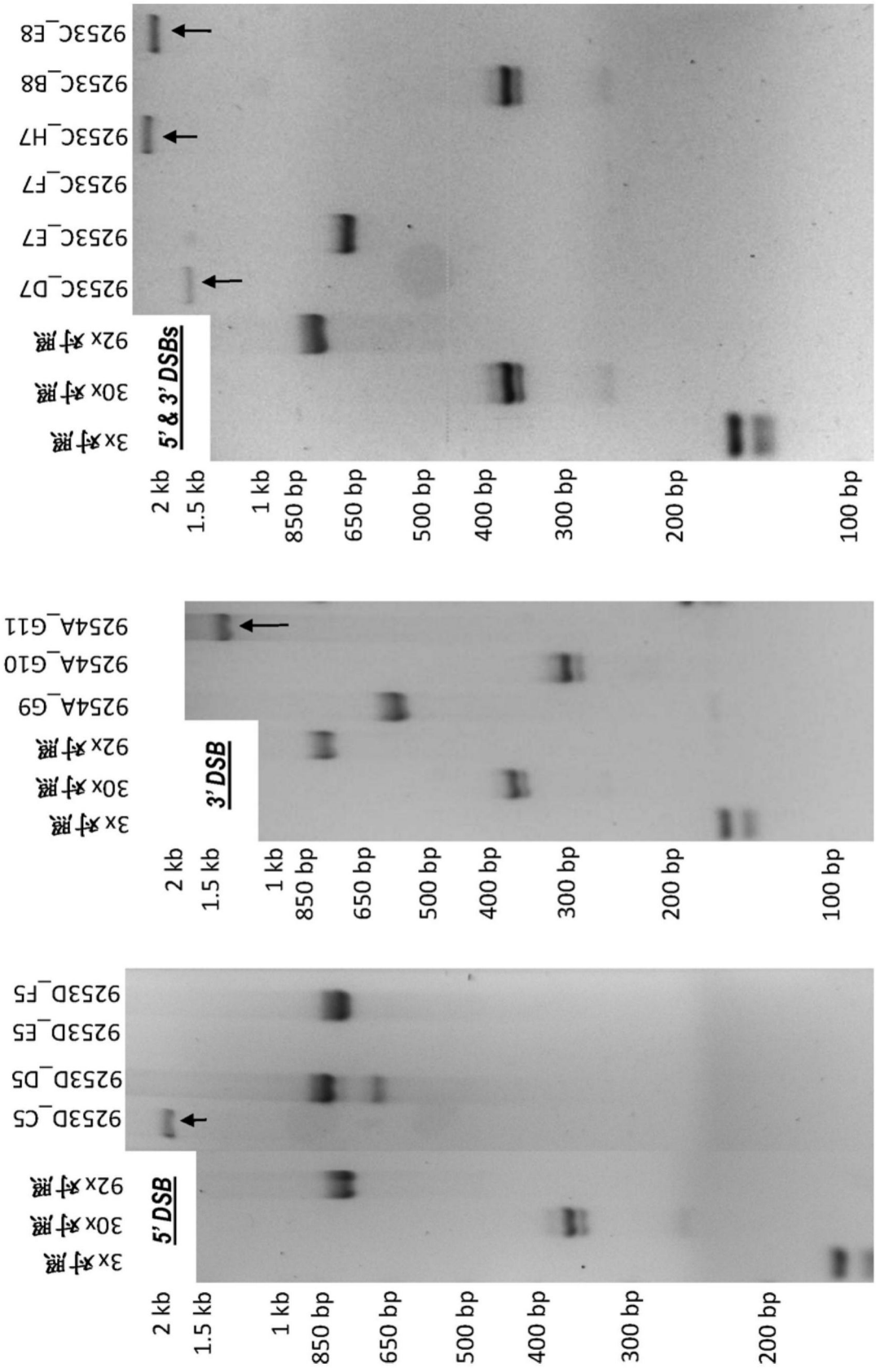

5' & 3' DSBs
3x 对照
30x 对照
92x 对照
9253C_D7
9253C_E7
9253C_F7
9253C_H7
9253C_B8
9253C_E8
2 kb
1.5 kb
1 kb
850 bp
650 bp
500 bp
400 bp
300 bp
200 bp
100 bp
3' DSB
3x 对照
30x 对照
92x 对照
9254A_G9
9254A_G10
9254A_G11
2 kb
1.5 kb
1 kb
850 bp
650 bp
500 bp
400 bp
300 bp
200 bp
100 bp
5' DSB
3x 对照
30x 对照
92x 对照
9253D_C5
9253D_D5
9253D_E5
9253D_F5
2 kb
1.5 kb
1 kb
850 bp
650 bp
500 bp
400 bp
300 bp
200 bp

图10

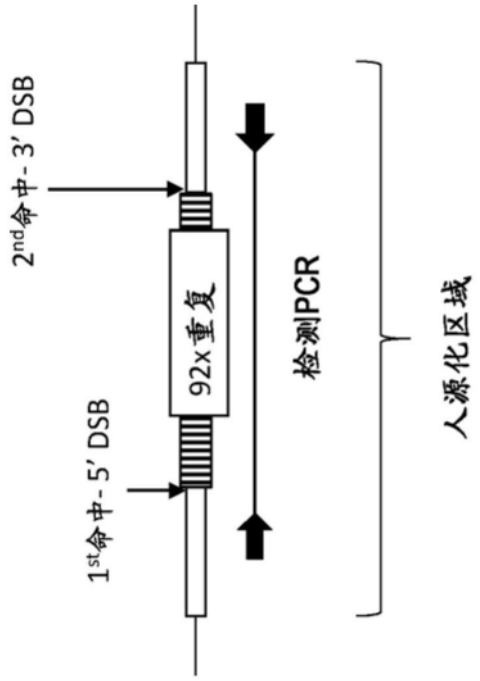


图11A

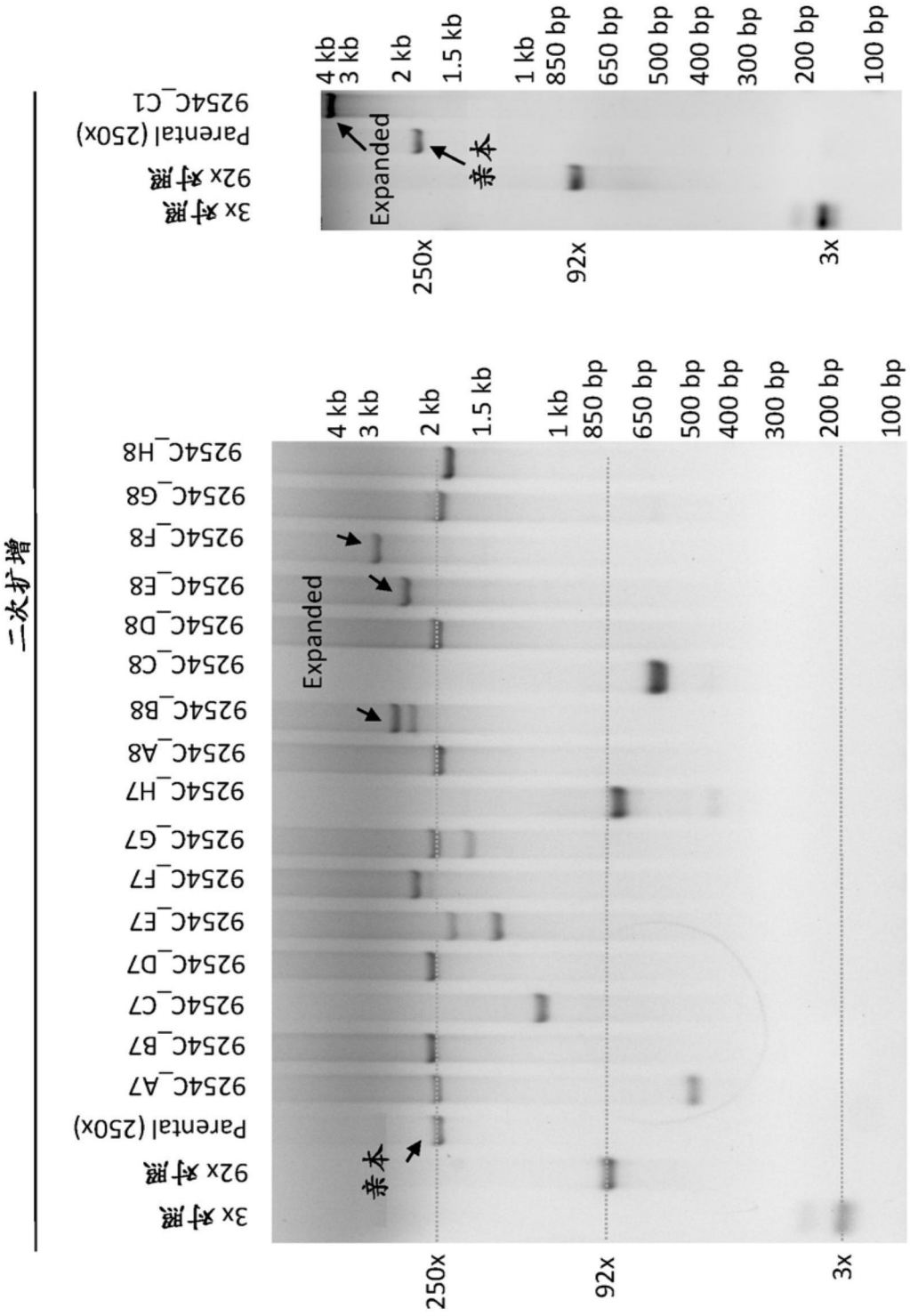

二次扩增
3x 对照
92x 对照
Parental (250x)
9254C_A7
9254C_B7
9254C_C7
9254C_D7
9254C_E7
9254C_F7
9254C_G7
9254C_H7
9254C_A8
9254C_B8
9254C_C8
9254C_D8
9254C_E8
9254C_F8
9254C_G8
9254C_H8
Expanded
亲本
4 kb
3 kb
2 kb
1.5 kb
1 kb
850 bp
650 bp
500 bp
400 bp
300 bp
200 bp
100 bp
250x
92x
3x
3x 对照
92x 对照
Parental (250x)
9254C_C1
Expanded
亲本
4 kb
3 kb
2 kb
1.5 kb
1 kb
850 bp
650 bp
500 bp
400 bp
300 bp
200 bp
100 bp
250x
92x
3x

图11B

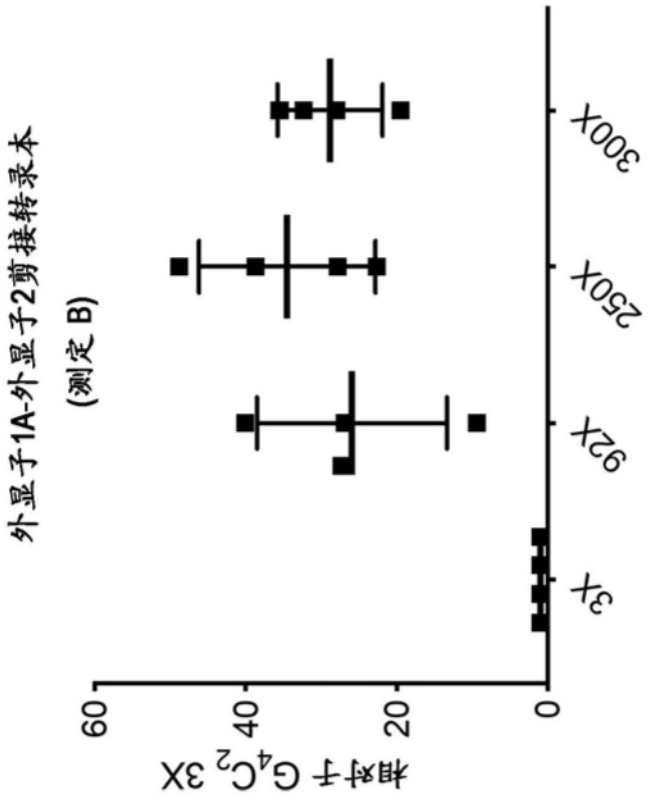


图12A

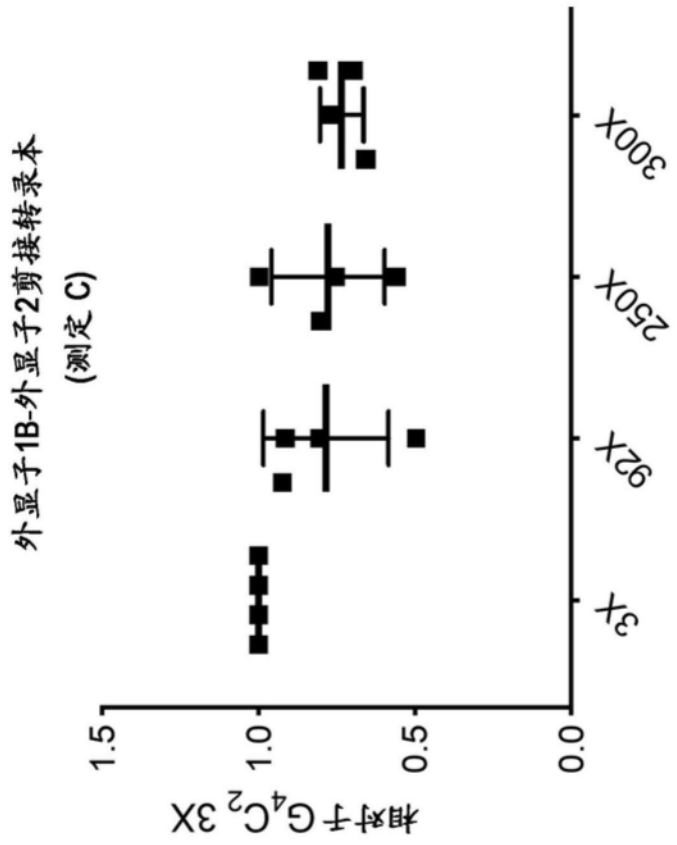

外显子1B-外显子2剪接转录本
(测定 C)
相对于 G_4C_2 3X
1.5
1.0
0.5
0.0
3x
92x
250x
300x

图12B

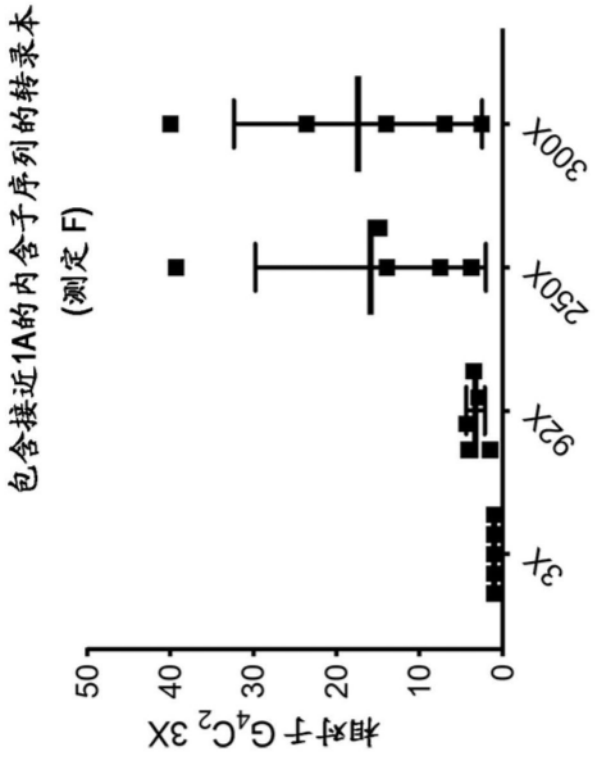

包含接近1A的内含子序列的转录本
(测定 F)
相对于 G_4C_2 3X
50
40
30
20
10
0
3x
92x
250x
300x

图12C

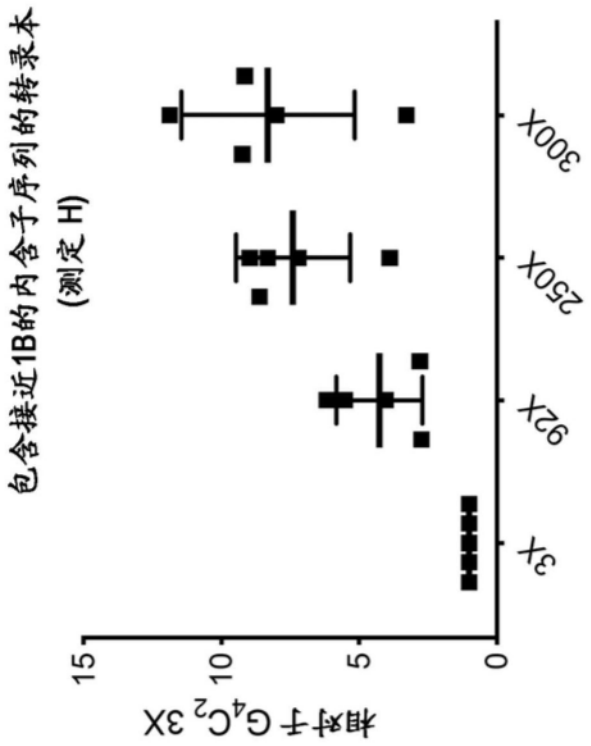

包含接近1B的内含子序列的转录本
(测定 H)
相对于 G_4C_2 3X
15
10
5
0
3x
92x
250x
300x

图12D

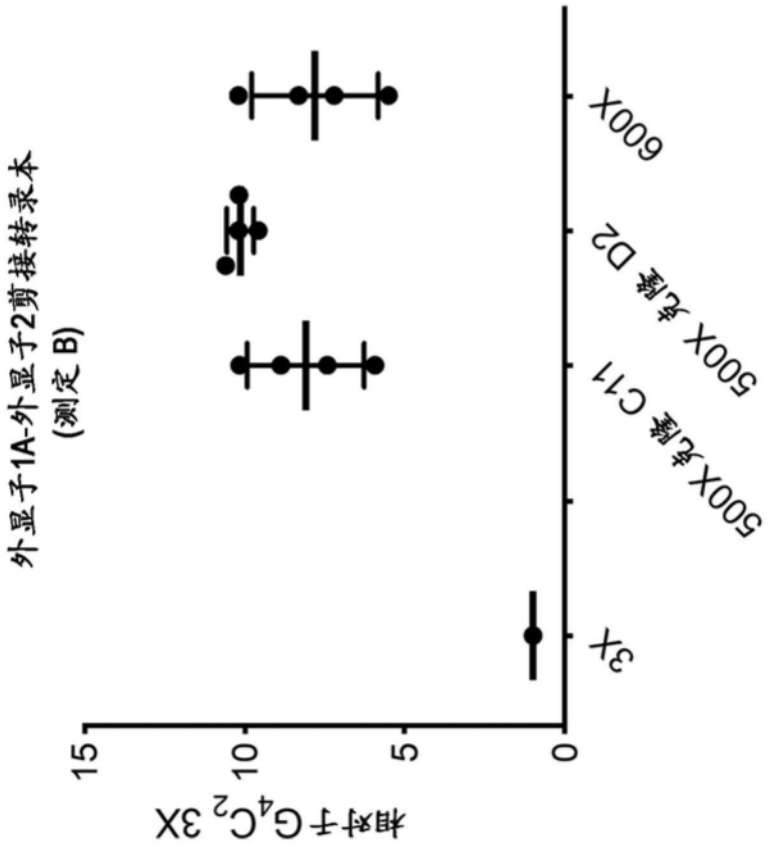

外显子1A-外显子2剪接转录本
(测定 B)
相对于 G_4C_2 3X
15
10
5
0
3x
500X 去除 C11
500X 去除 D2
600X

图12E

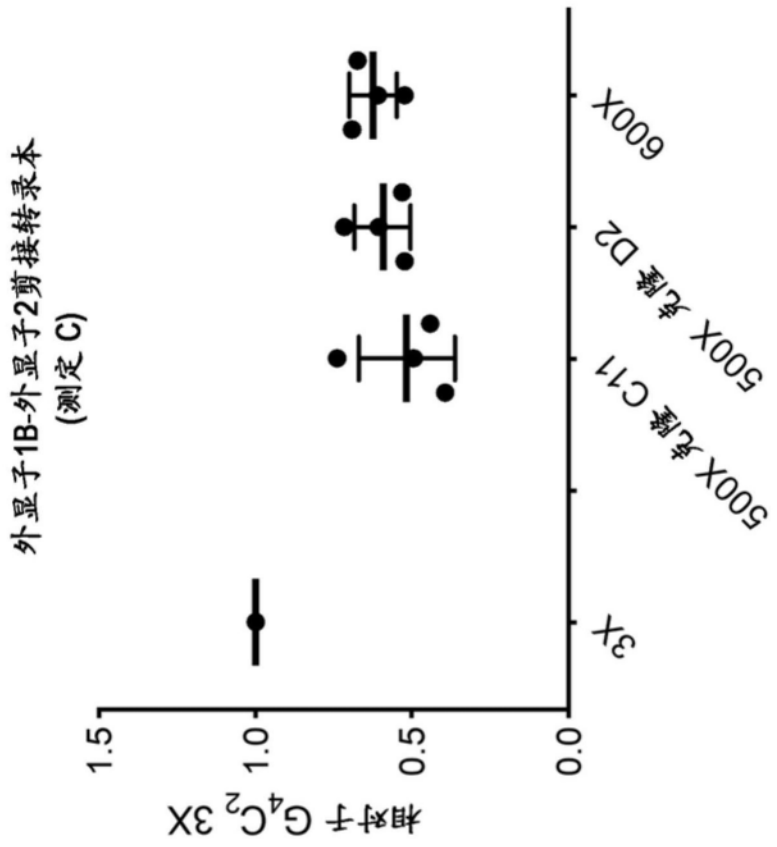

外显子1B-外显子2剪接转录本
(测定 C)
相对于 G_4C_2 3X
1.5
1.0
0.5
0.0
3X
500X 长度 C11
500X 长度 D2
600X

图12F

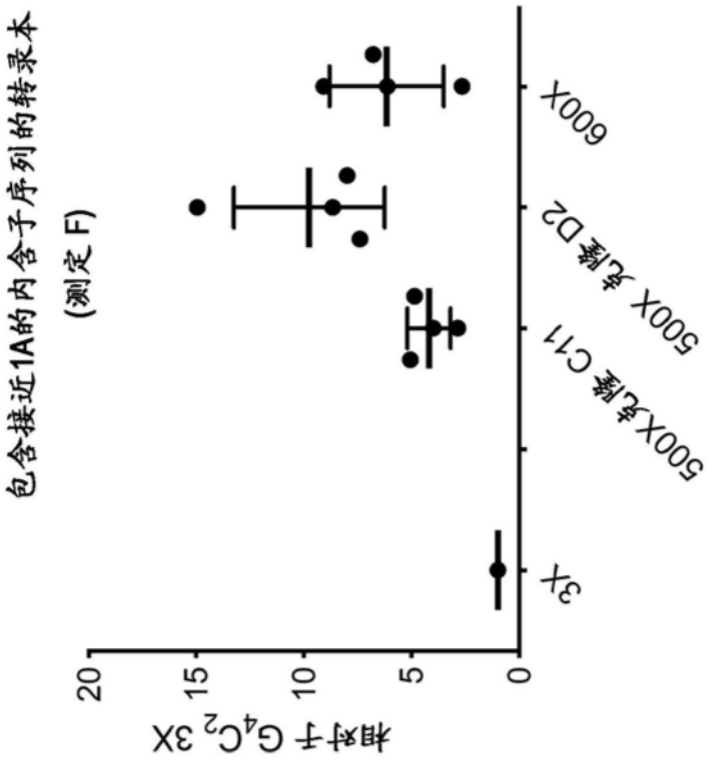

包含接近1A的内含子序列的转录本
(测定 F)
相对于 G_4C_2 3X
20
15
10
5
0
3X
500X 先降 C11
500X 先降 D2
600X

图12G

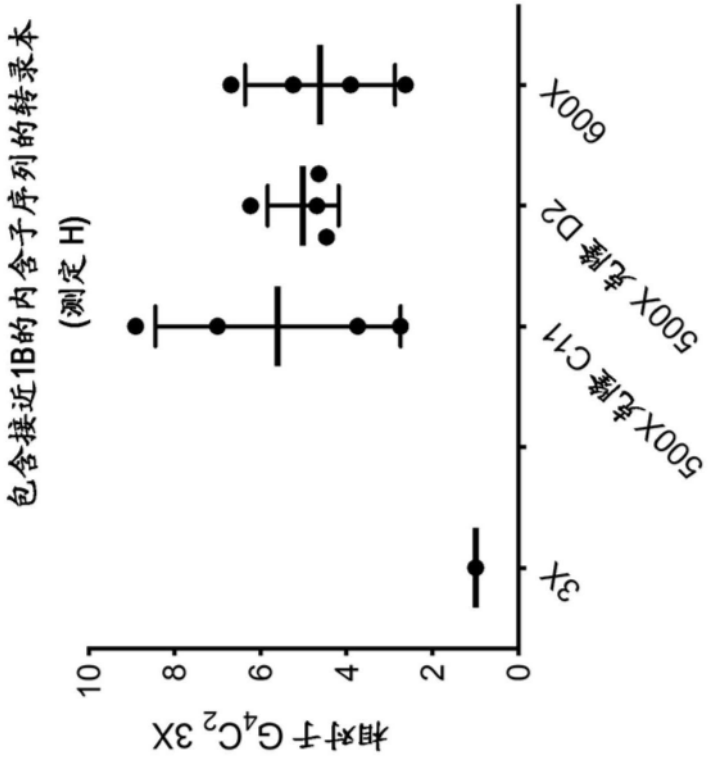

包含接近1B的内含子序列的转录本
(测定 H)
相对于 G_4C_2 3X
10
8
6
4
2
0
3X
500X 先降 C11
500X 先降 D2
600X

图12H

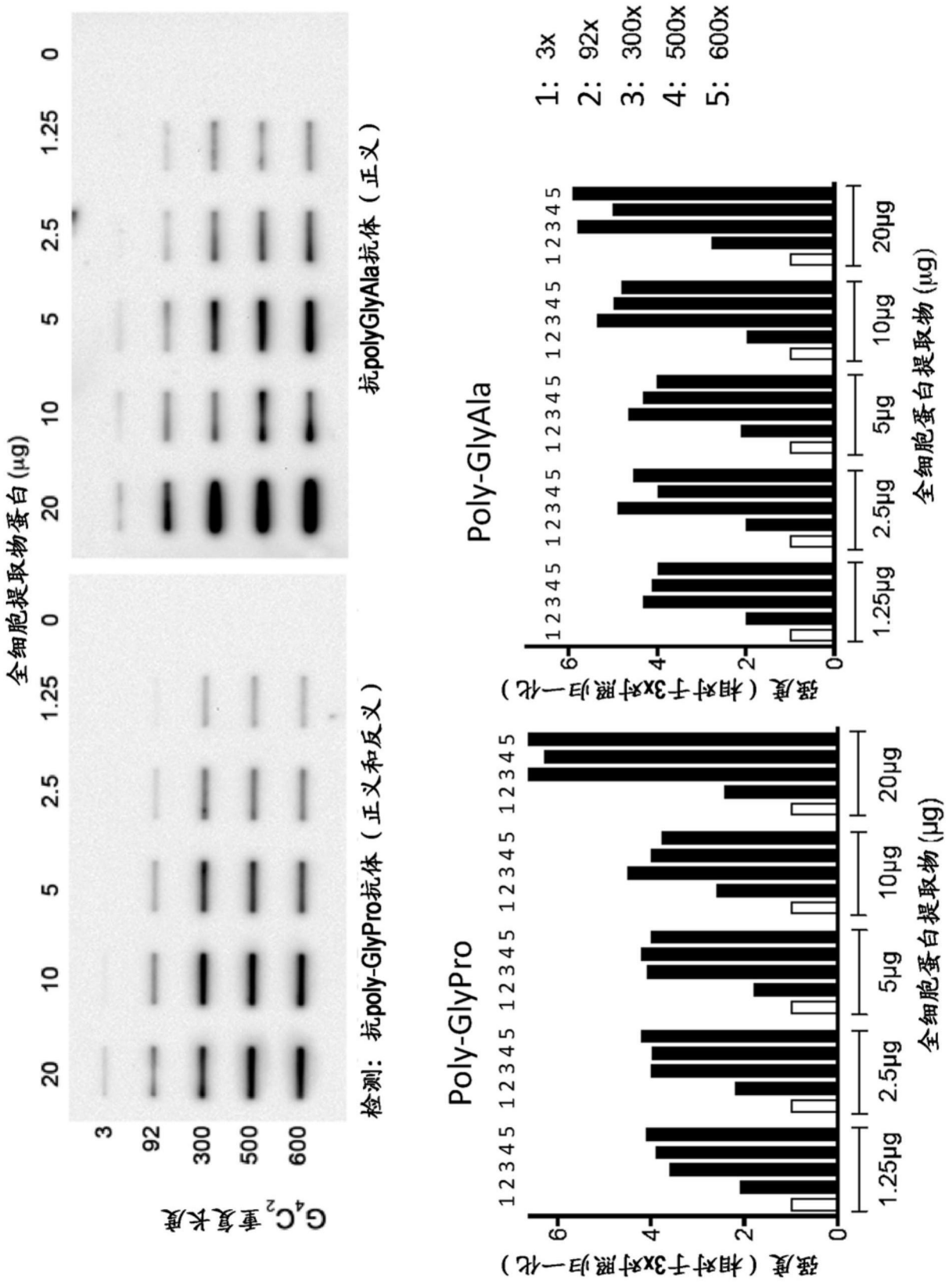

检测：抗poly-GlyPro抗体（正义和反义）
抗polyGlyAla抗体（正义）
全细胞提取物蛋白(μg)
20 10 5 2.5 1.25 0
3
92
300
500
600
G_4C_2重复序列
Poly-GlyPro
Poly-GlyAla
1: 3x
2: 92x
3: 300x
4: 500x
5: 600x
1 2 3 4 5
1.25μg 2.5μg 5μg 10μg 20μg
全细胞蛋白提取物 (μg)
强度（相对于3x对照归一化）
0 2 4 6

图13

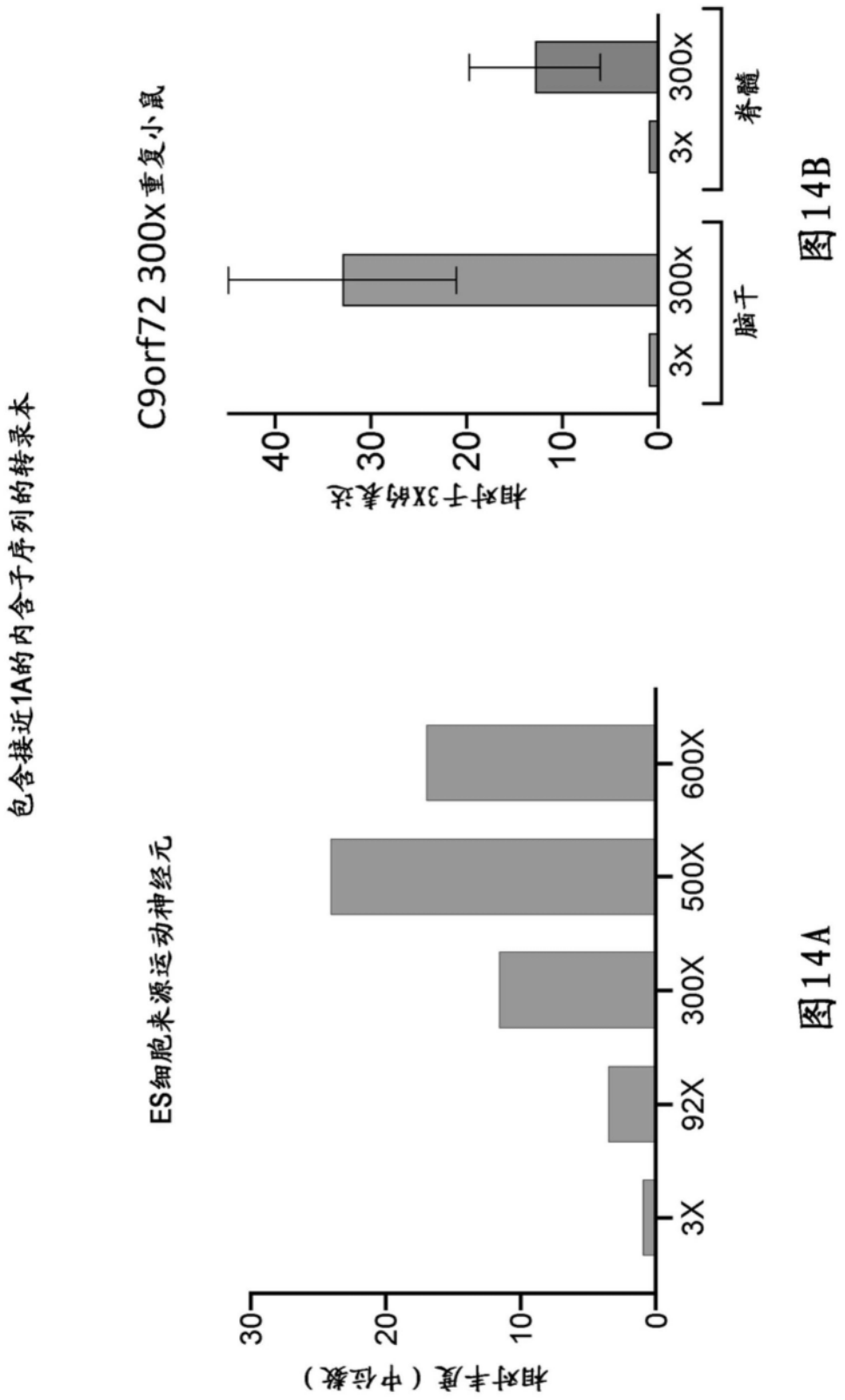

包含接近1A的内含子序列的转录本
ES细胞来源运动神经元
相对丰度（中位数）
0
10
20
30
3X
92X
300X
500X
600X
C9orf72 300x重复小鼠
相对于3X的表达
0
10
20
30
40
3x
300x
脑干
3x
300x
脊髓

图14A

图14B

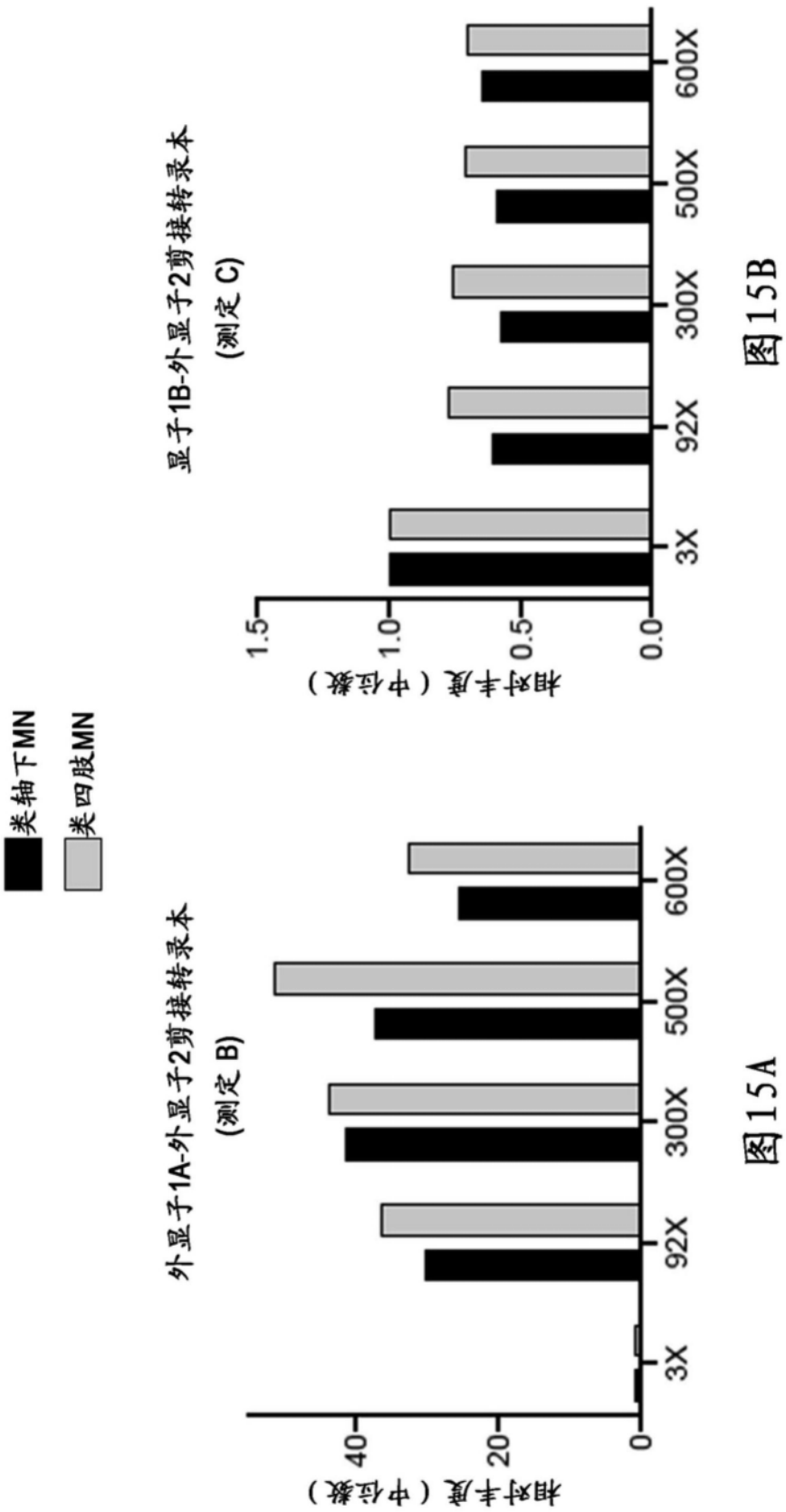

类轴下MN
类四肢MN
外显子1A-外显子2剪接转录本
(测定 B)
相对丰度（中位数）
0
20
40
3X
92X
300X
500X
600X
显子1B-外显子2剪接转录本
(测定 C)
相对丰度（中位数）
0.0
0.5
1.0
1.5
3X
92X
300X
500X
600X

图15A

图15B

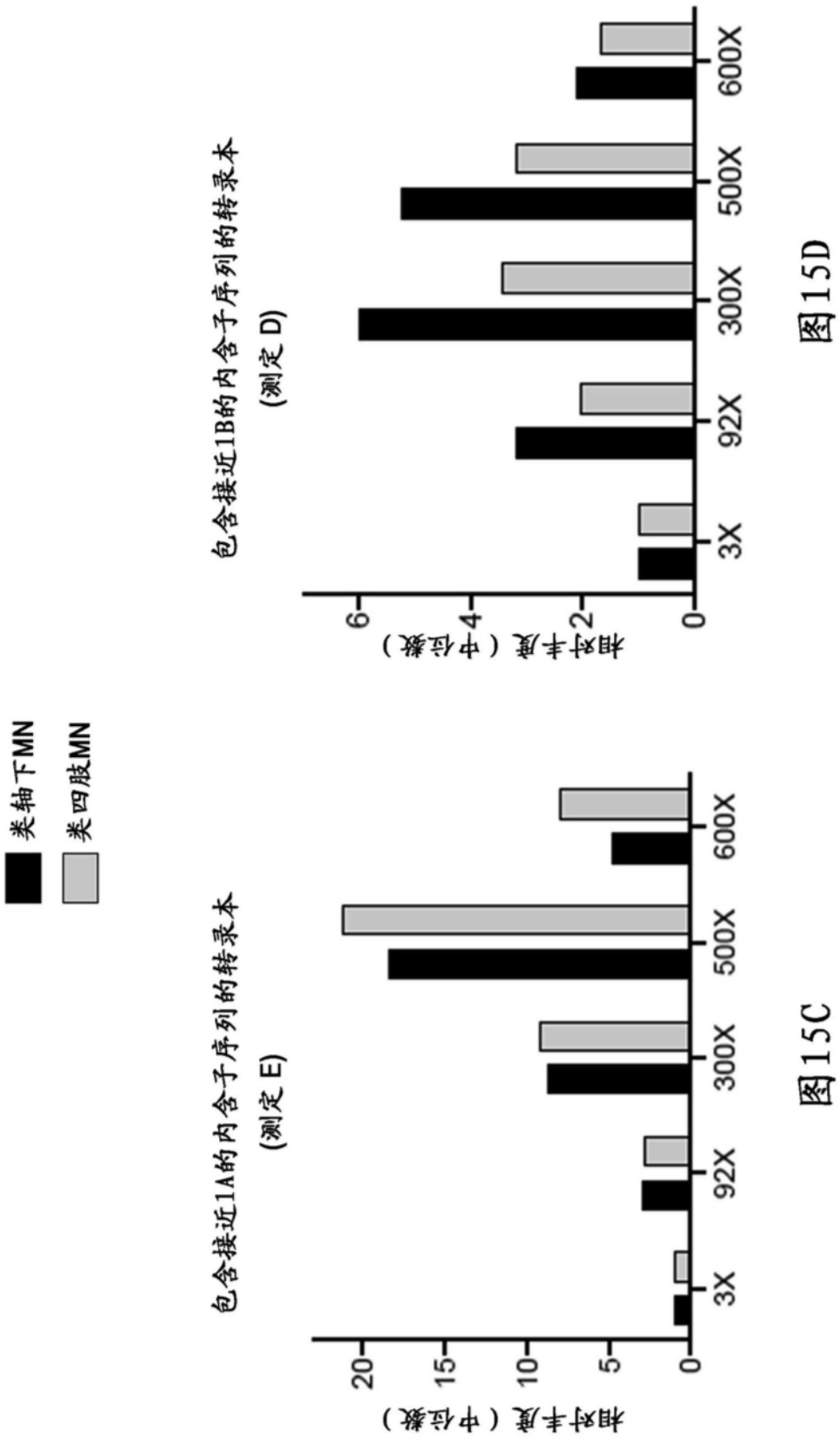

类轴下MN
类四肢MN
包含接近1A的内含子序列的转录本
(测定 E)
相对丰度（中位数）
0
5
10
15
20
3X
92X
300X
500X
600X
包含接近1B的内含子序列的转录本
(测定 D)
相对丰度（中位数）
0
2
4
6
3X
92X
300X
500X
600X

图15C

图15D

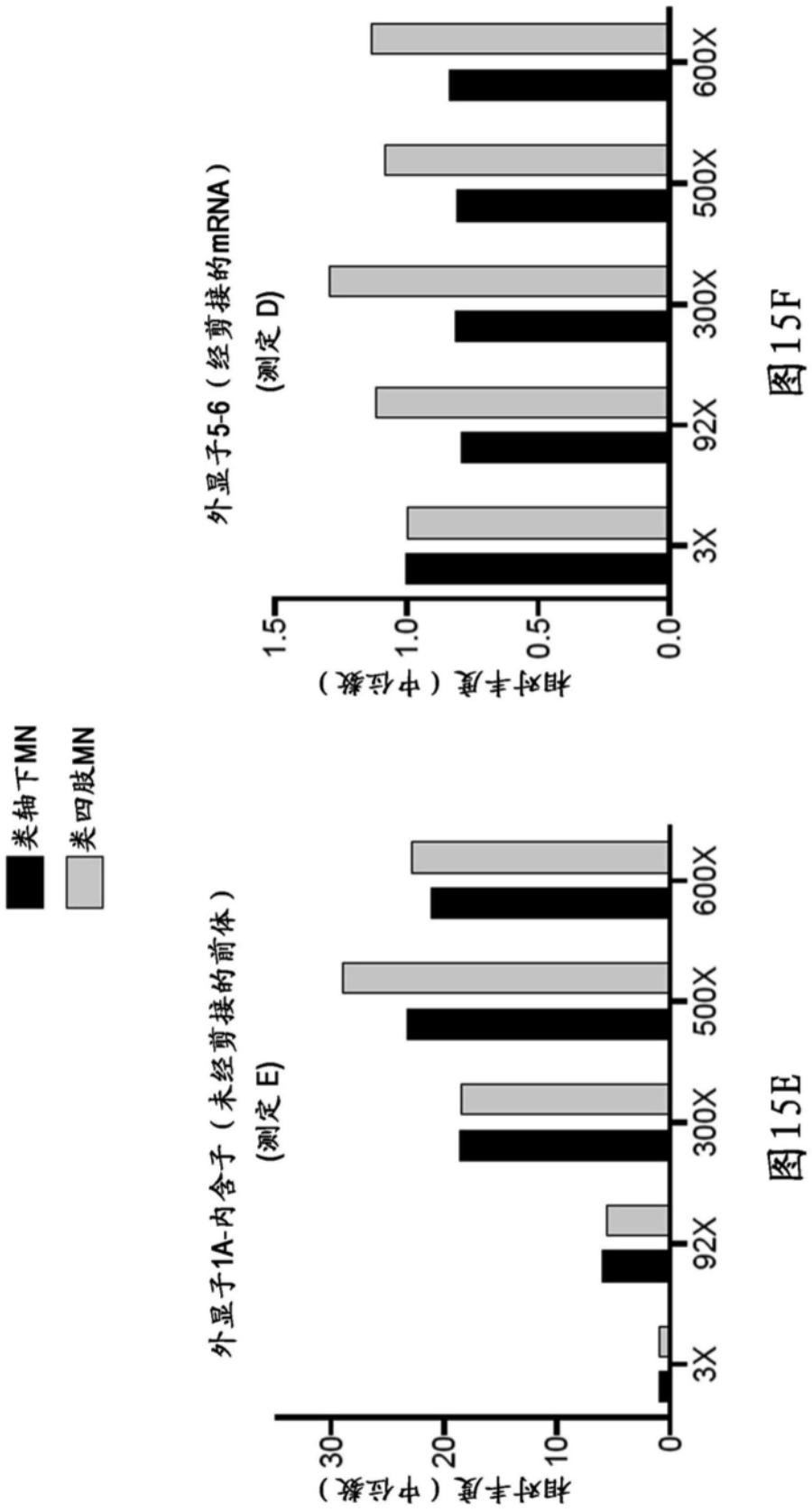

外显子1A-内含子（未经剪接的前体）
(测定 E)
相对丰度（中位数）
0
10
20
30
3X
92X
300X
500X
600X
外显子5-6（经剪接的mRNA）
(测定 D)
相对丰度（中位数）
0.0
0.5
1.0
1.5
3X
92X
300X
500X
600X
类轴下MN
类四肢MN

图15E

图15F

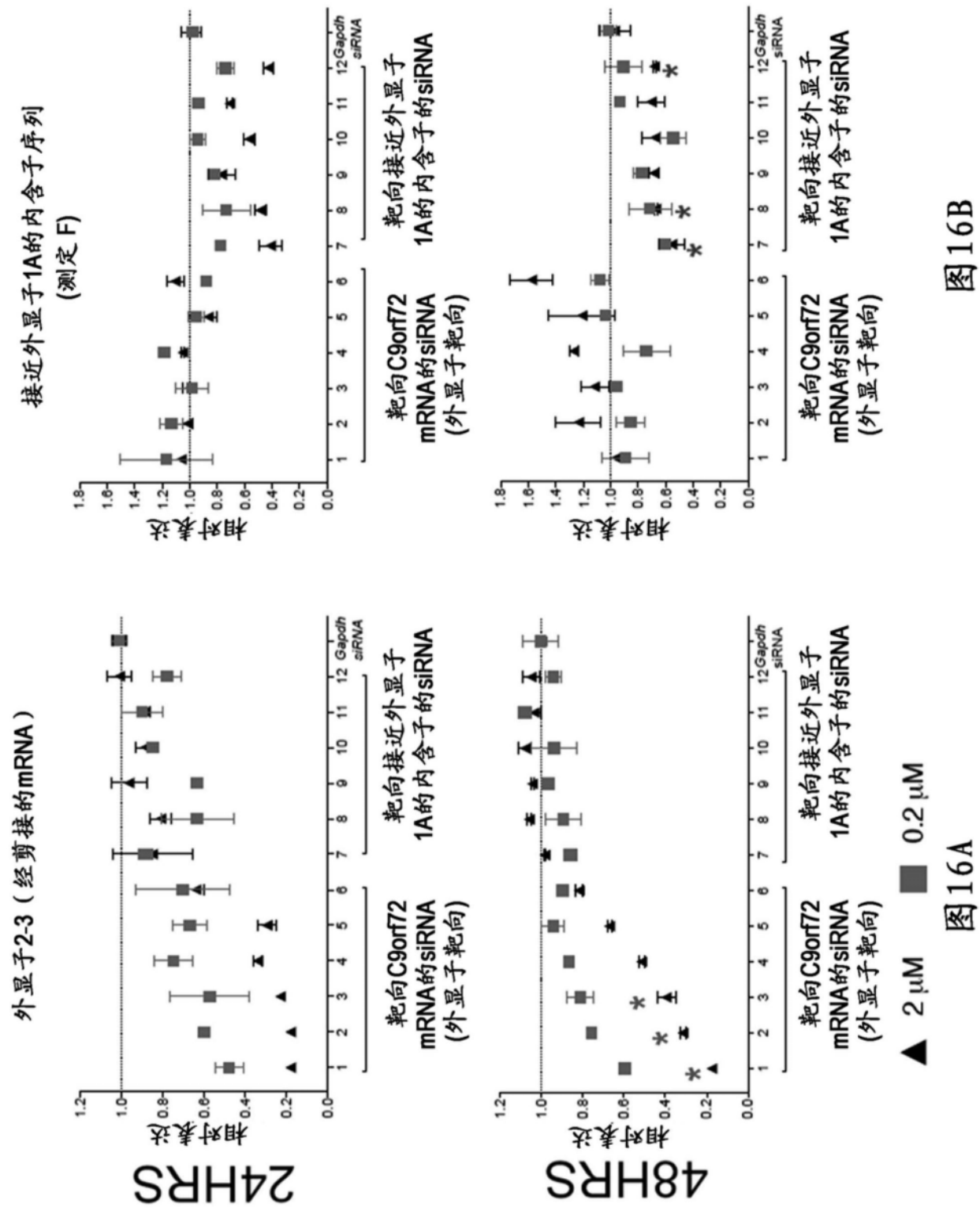

外显子2-3（经剪接的mRNA）
接近外显子1A的内含子序列
(测定 F)
24HRS
48HRS
相对表达
靶向C9orf72
mRNA的siRNA
(外显子靶向)
靶向接近外显子
1A的内含子的siRNA
Gapdh
siRNA
2 μM
0.2 μM

图16A

图16B

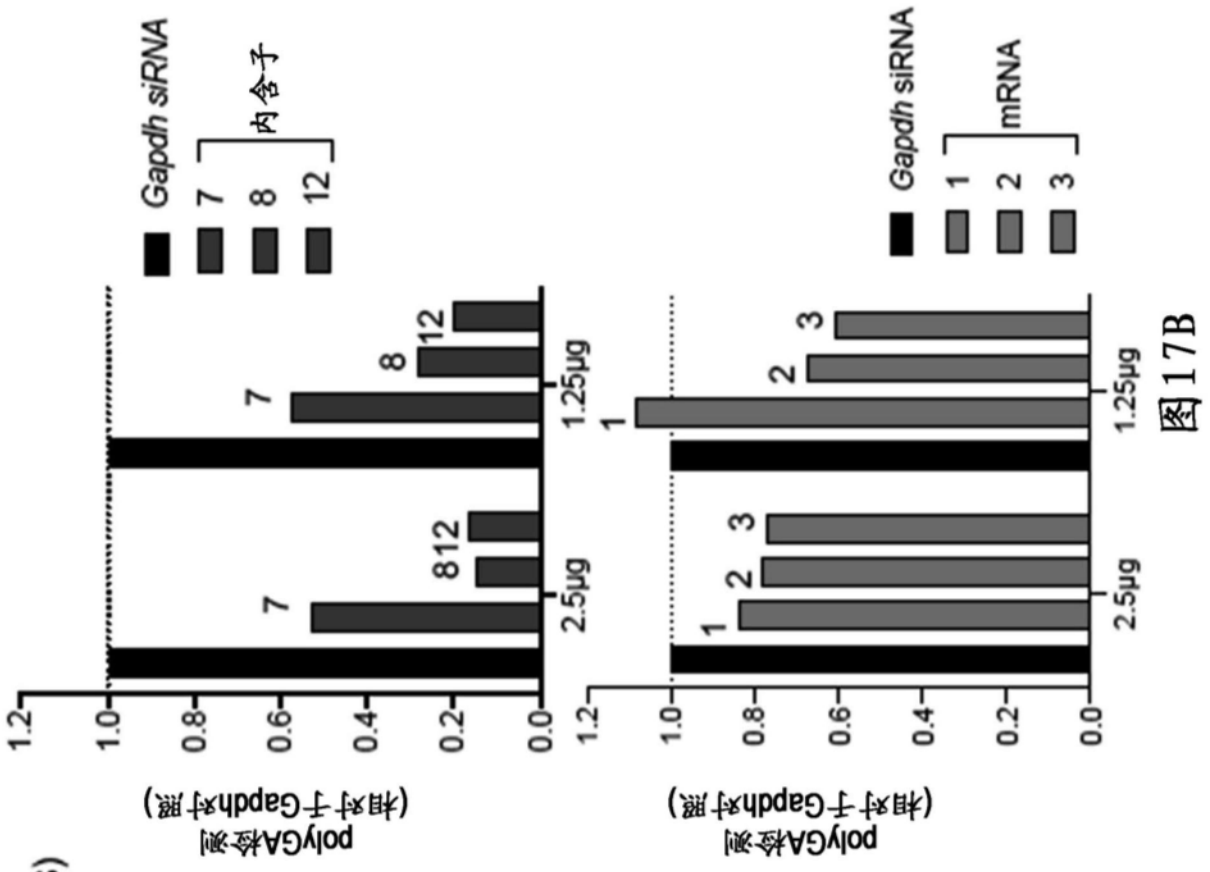

Gapdh siRNA
内含子
7
8
12
mRNA
1
2
3
polyGA检测
(相对于Gapdh对照)
1.25μg
2.5μg

图17B

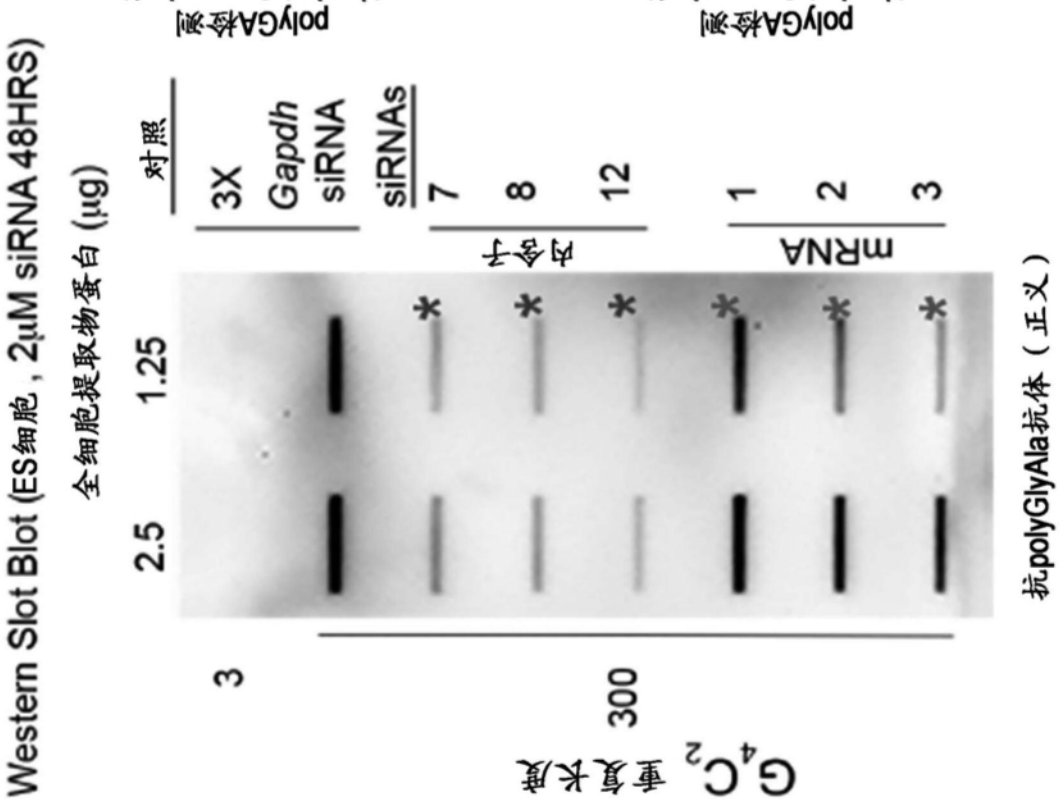

Western Slot Blot (ES细胞，2μM siRNA 48HRS)
全细胞提取物蛋白 (μg)
2.5
1.25
对照
3X Gapdh siRNA
siRNAs
内含子
mRNA
抗polyGlyAla抗体（正义）
G4C2 重复长度
300
3

图17A

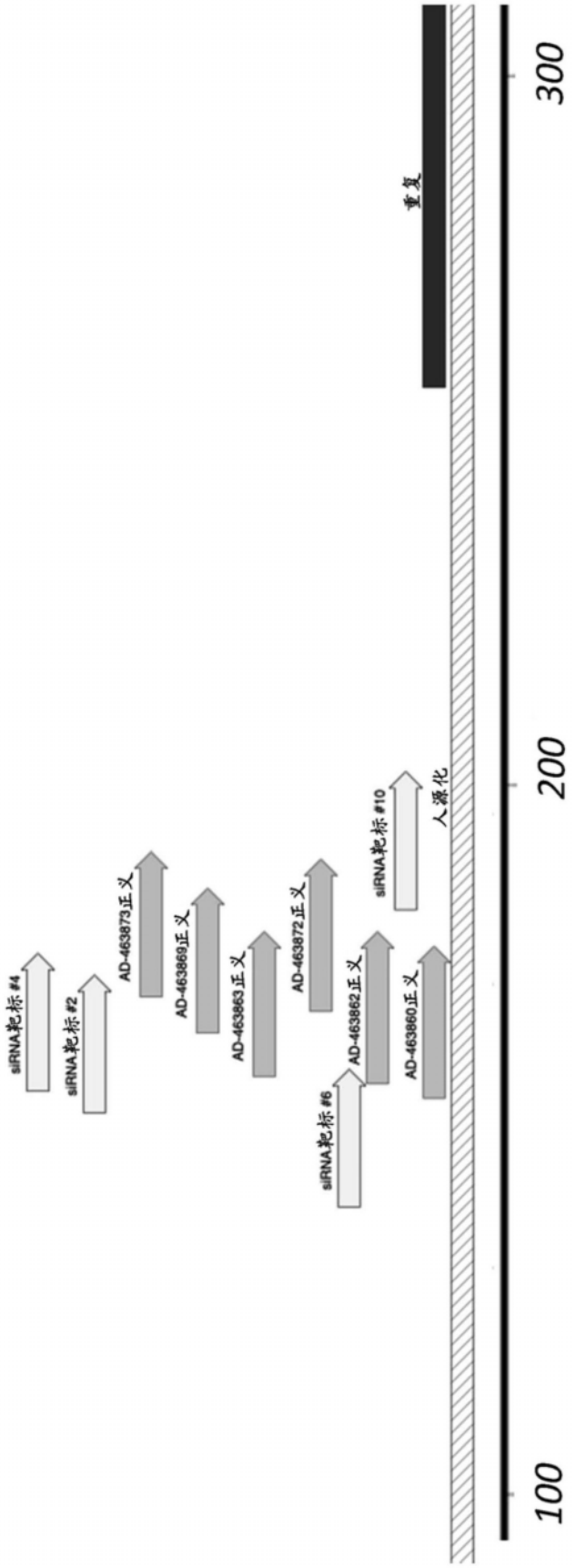

重复
人源化
100
200
300
siRNA靶标#4
siRNA靶标#2
AD-463873正义
AD-463869正义
AD-463863正义
AD-463872正义
AD-463862正义
AD-463860正义
siRNA靶标#6
siRNA靶标#10

图18A

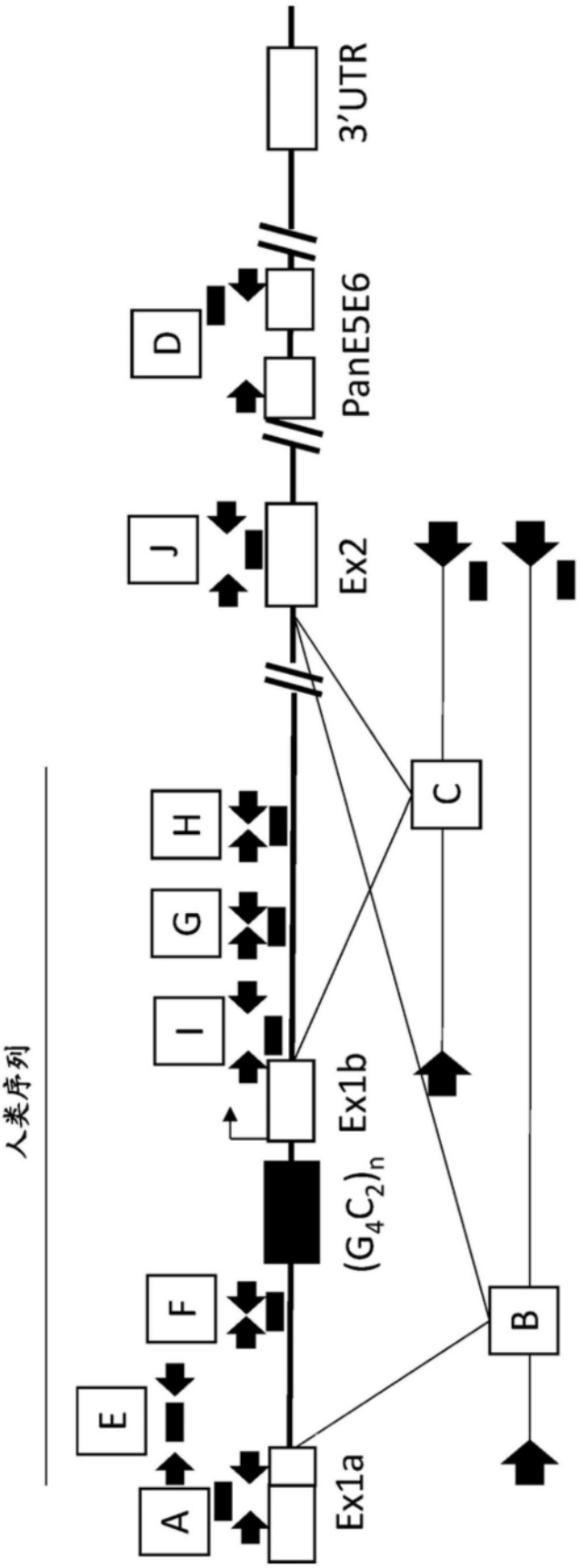


图18B

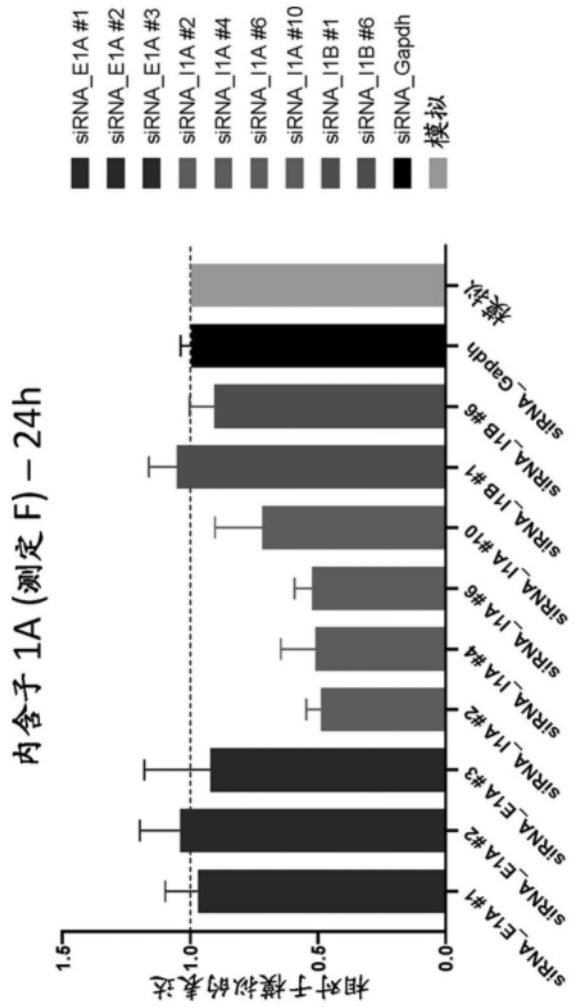

siRNA_E1A #1
siRNA_E1A #2
siRNA_E1A #3
siRNA_I1A #2
siRNA_I1A #4
siRNA_I1A #6
siRNA_I1A #10
siRNA_I1B #1
siRNA_I1B #6
siRNA_Gapdh
模拟
内含子 1A (测定 F) – 24h
相对于模拟的表达
1.5
1.0
0.5
0.0
siRNA_E1A #1
siRNA_E1A #2
siRNA_E1A #3
siRNA_I1A #2
siRNA_I1A #4
siRNA_I1A #6
siRNA_I1A #10
siRNA_I1B #1
siRNA_I1B #6
siRNA_Gapdh
模拟

图19A

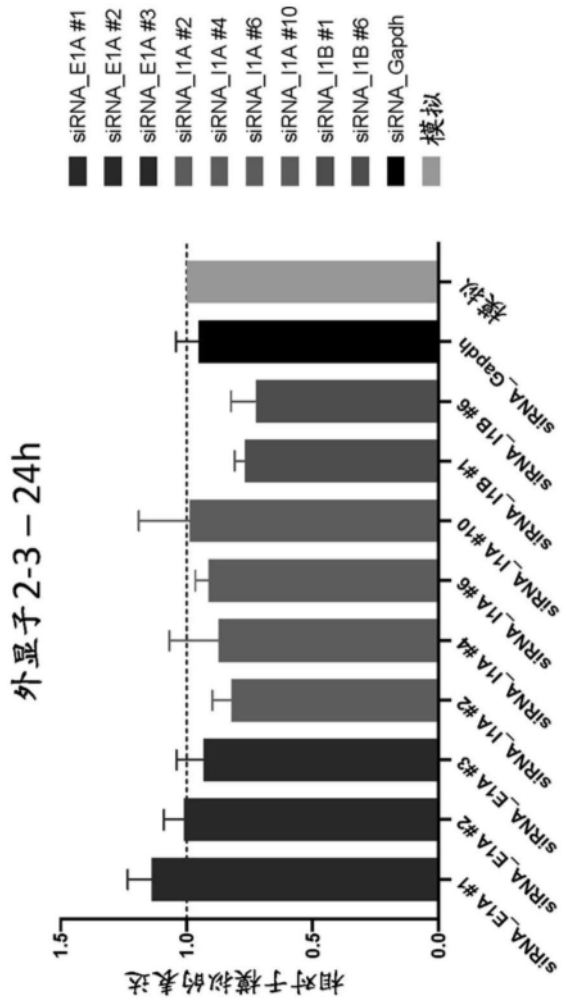

siRNA_E1A #1
siRNA_E1A #2
siRNA_E1A #3
siRNA_I1A #2
siRNA_I1A #4
siRNA_I1A #6
siRNA_I1A #10
siRNA_I1B #1
siRNA_I1B #6
siRNA_Gapdh
模拟
外显子2-3－24h
相对于模拟的表达
1.5
1.0
0.5
0.0
siRNA_E1A #1
siRNA_E1A #2
siRNA_E1A #3
siRNA_I1A #2
siRNA_I1A #4
siRNA_I1A #6
siRNA_I1A #10
siRNA_I1B #1
siRNA_I1B #6
siRNA_Gapdh
模拟

图19B

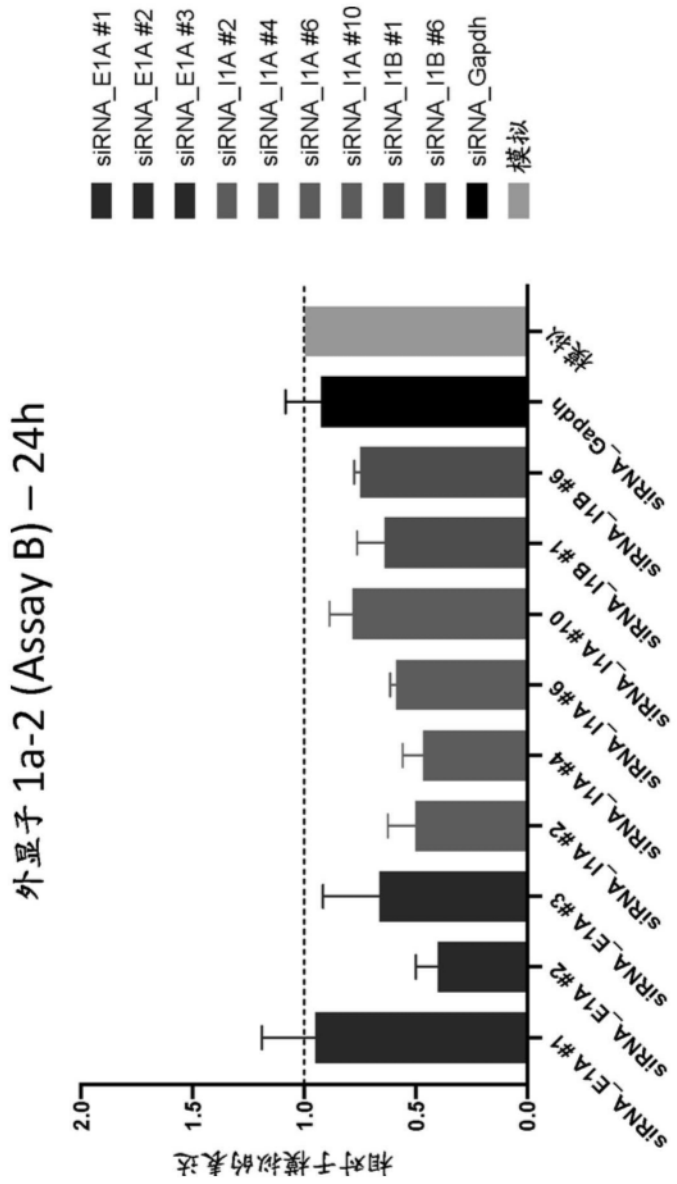

siRNA_E1A #1
siRNA_E1A #2
siRNA_E1A #3
siRNA_I1A #2
siRNA_I1A #4
siRNA_I1A #6
siRNA_I1A #10
siRNA_I1B #1
siRNA_I1B #6
siRNA_Gapdh
模拟
外显子 1a-2 (Assay B) – 24h
相对于模拟的表达
2.0
1.5
1.0
0.5
0.0
siRNA_E1A #1
siRNA_E1A #2
siRNA_E1A #3
siRNA_I1A #2
siRNA_I1A #4
siRNA_I1A #6
siRNA_I1A #10
siRNA_I1B #1
siRNA_I1B #6
siRNA_Gapdh
模拟

图19C

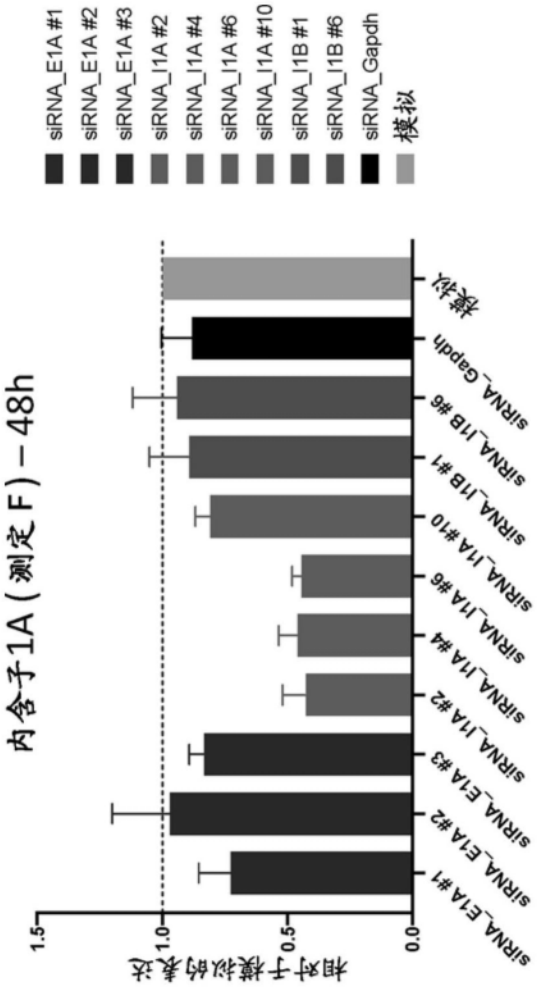

siRNA_E1A #1
siRNA_E1A #2
siRNA_E1A #3
siRNA_I1A #2
siRNA_I1A #4
siRNA_I1A #6
siRNA_I1A #10
siRNA_I1B #1
siRNA_I1B #6
siRNA_Gapdh
模拟
内含子1A（测定F）– 48h
相对于模拟的表达
1.5
1.0
0.5
0.0
siRNA_E1A #1
siRNA_E1A #2
siRNA_E1A #3
siRNA_I1A #2
siRNA_I1A #4
siRNA_I1A #6
siRNA_I1A #10
siRNA_I1B #1
siRNA_I1B #6
siRNA_Gapdh
模拟

图20A

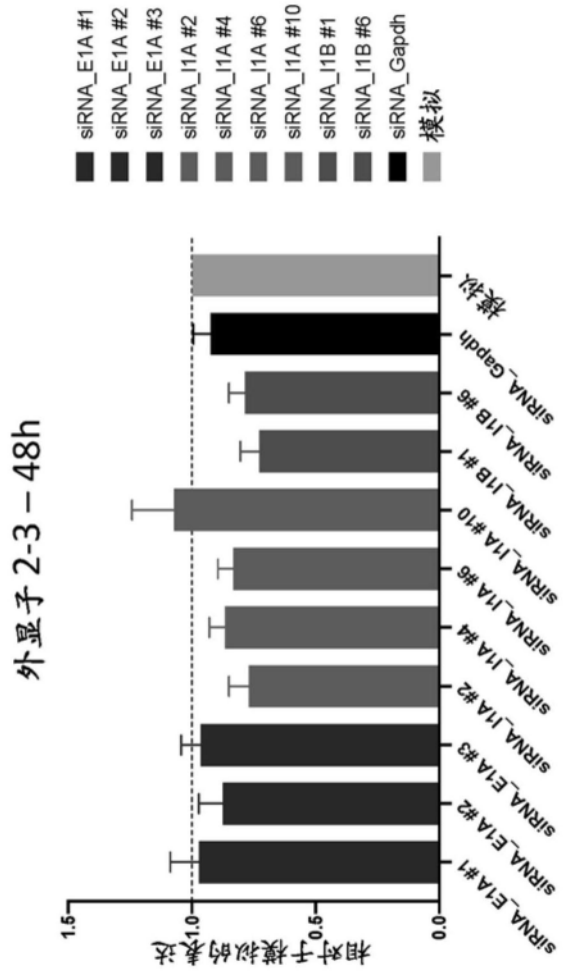

siRNA_E1A #1
siRNA_E1A #2
siRNA_E1A #3
siRNA_I1A #2
siRNA_I1A #4
siRNA_I1A #6
siRNA_I1A #10
siRNA_I1B #1
siRNA_I1B #6
siRNA_Gapdh
模拟
外显子2-3－48h
相对于模拟的表达
1.5
1.0
0.5
0.0
siRNA_E1A #1
siRNA_E1A #2
siRNA_E1A #3
siRNA_I1A #2
siRNA_I1A #4
siRNA_I1A #6
siRNA_I1A #10
siRNA_I1B #1
siRNA_I1B #6
siRNA_Gapdh
模拟

图20B

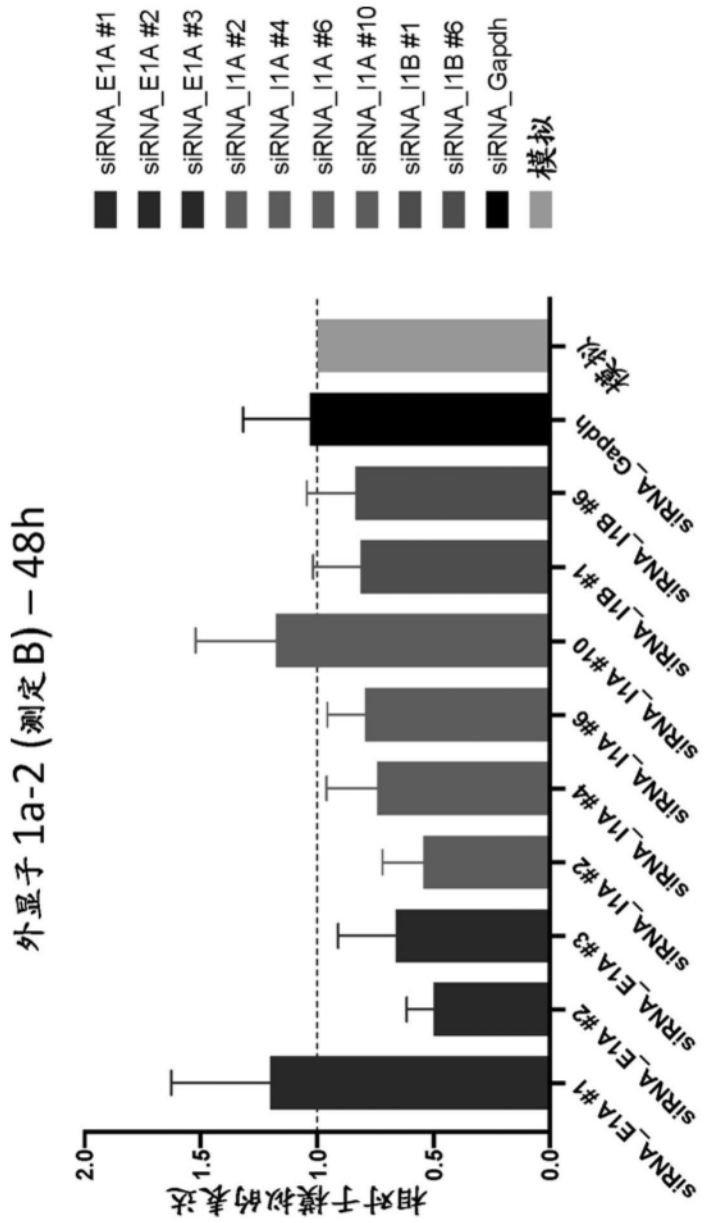

siRNA_E1A #1
siRNA_E1A #2
siRNA_E1A #3
siRNA_I1A #2
siRNA_I1A #4
siRNA_I1A #6
siRNA_I1A #10
siRNA_I1B #1
siRNA_I1B #6
siRNA_Gapdh
模拟
外显子1a-2 (测定B) – 48h
相对于模拟的表达
2.0
1.5
1.0
0.5
0.0
siRNA_E1A #1
siRNA_E1A #2
siRNA_E1A #3
siRNA_I1A #2
siRNA_I1A #4
siRNA_I1A #6
siRNA_I1A #10
siRNA_I1B #1
siRNA_I1B #6
siRNA_Gapdh
模拟

图20C

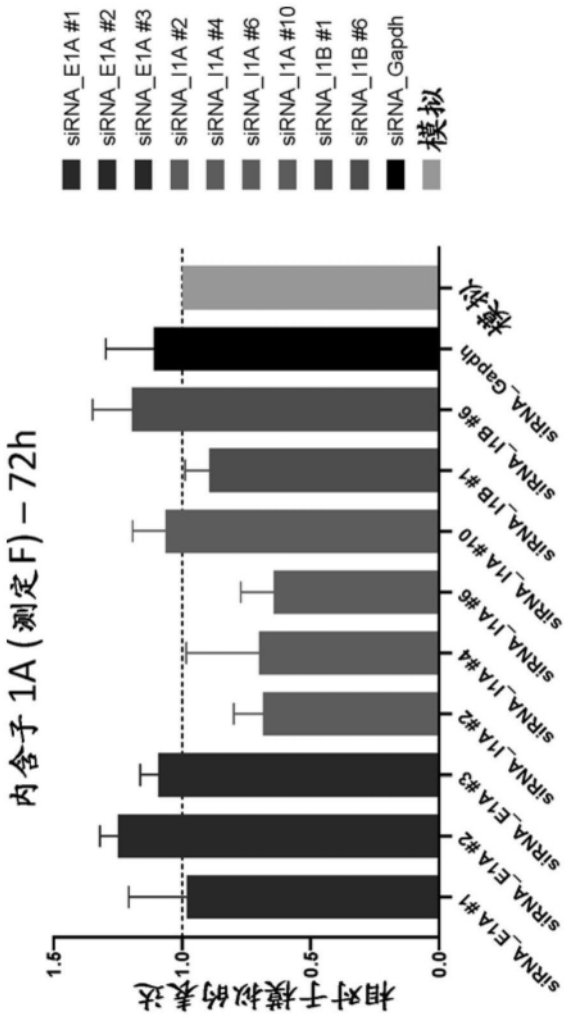

siRNA_E1A #1
siRNA_E1A #2
siRNA_E1A #3
siRNA_I1A #2
siRNA_I1A #4
siRNA_I1A #6
siRNA_I1A #10
siRNA_I1B #1
siRNA_I1B #6
siRNA_Gapdh
模拟
内含子 1A (测定 F) – 72h
相对于模拟的表达
1.5
1.0
0.5
0.0
siRNA_E1A #1
siRNA_E1A #2
siRNA_E1A #3
siRNA_I1A #2
siRNA_I1A #4
siRNA_I1A #6
siRNA_I1A #10
siRNA_I1B #1
siRNA_I1B #6
siRNA_Gapdh
模拟

图21A

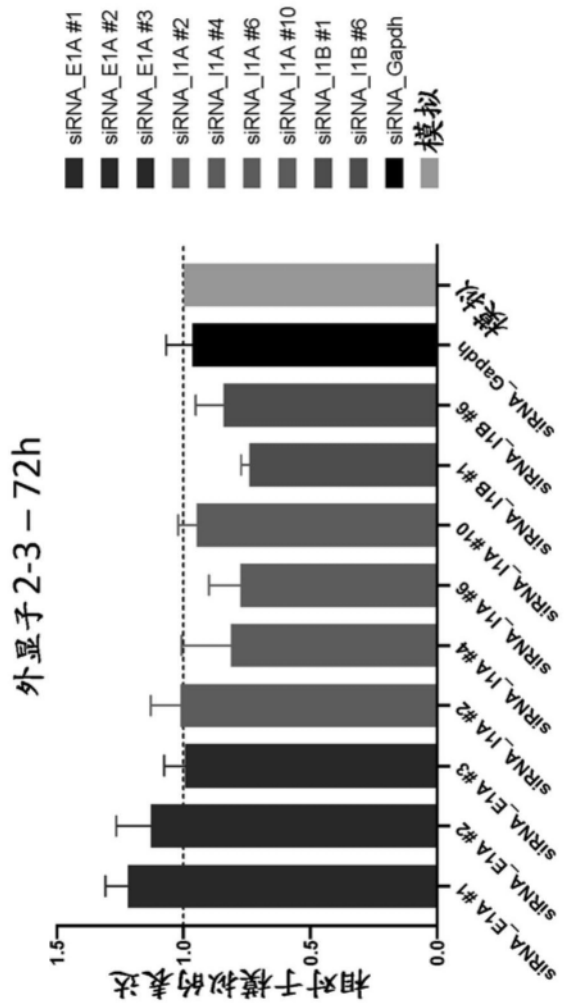

siRNA_E1A #1
siRNA_E1A #2
siRNA_E1A #3
siRNA_I1A #2
siRNA_I1A #4
siRNA_I1A #6
siRNA_I1A #10
siRNA_I1B #1
siRNA_I1B #6
siRNA_Gapdh
模拟
外显子2-3－72h
相对于模拟的表达
1.5
1.0
0.5
0.0
siRNA_E1A #1
siRNA_E1A #2
siRNA_E1A #3
siRNA_I1A #2
siRNA_I1A #4
siRNA_I1A #6
siRNA_I1A #10
siRNA_I1B #1
siRNA_I1B #6
siRNA_Gapdh
模拟

图21B

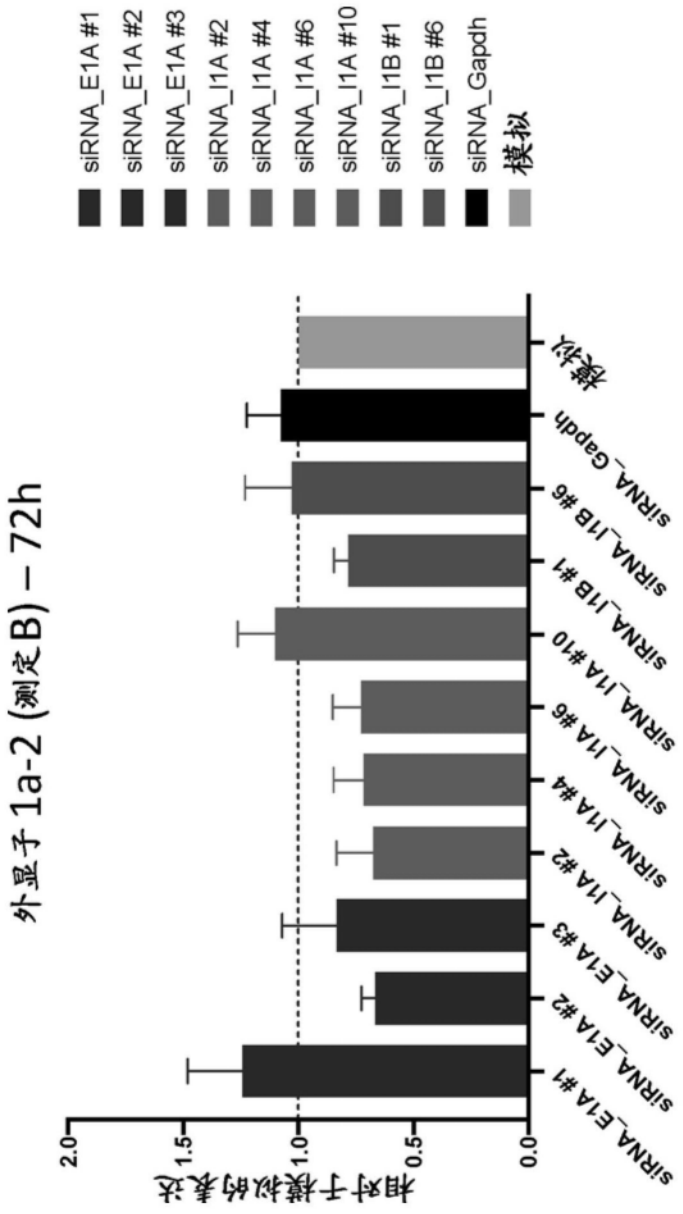

siRNA_E1A #1
siRNA_E1A #2
siRNA_E1A #3
siRNA_I1A #2
siRNA_I1A #4
siRNA_I1A #6
siRNA_I1A #10
siRNA_I1B #1
siRNA_I1B #6
siRNA_Gapdh
模拟
外显子 1a-2 (测定B) – 72h
相对于模拟的表达
2.0
1.5
1.0
0.5
0.0
siRNA_E1A #1
siRNA_E1A #2
siRNA_E1A #3
siRNA_I1A #2
siRNA_I1A #4
siRNA_I1A #6
siRNA_I1A #10
siRNA_I1B #1
siRNA_I1B #6
siRNA_Gapdh
模拟

图21C

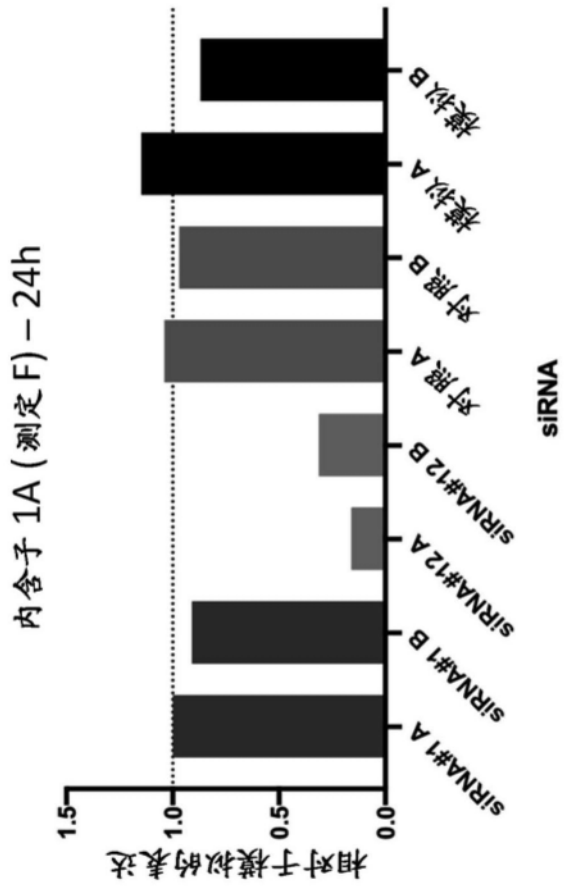

内含子 1A (测定 F) – 24h
相对于模拟的表达
0.0
0.5
1.0
1.5
siRNA
模拟 B
模拟 A
对照 B
对照 A
siRNA#12 B
siRNA#12 A
siRNA#1 B
siRNA#1 A

图22A

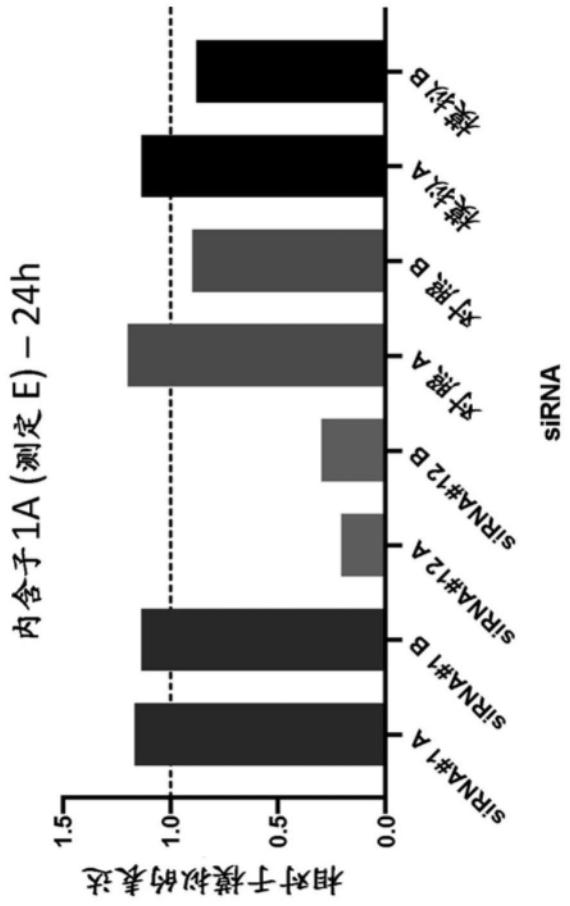

相对于模拟的表达
1.5
1.0
0.5
0.0
siRNA#1 A
siRNA#1 B
siRNA#12 A
siRNA#12 B
对照 A
对照 B
模拟 A
模拟 B
siRNA
内含子1A (测定 E) – 24h

图22B

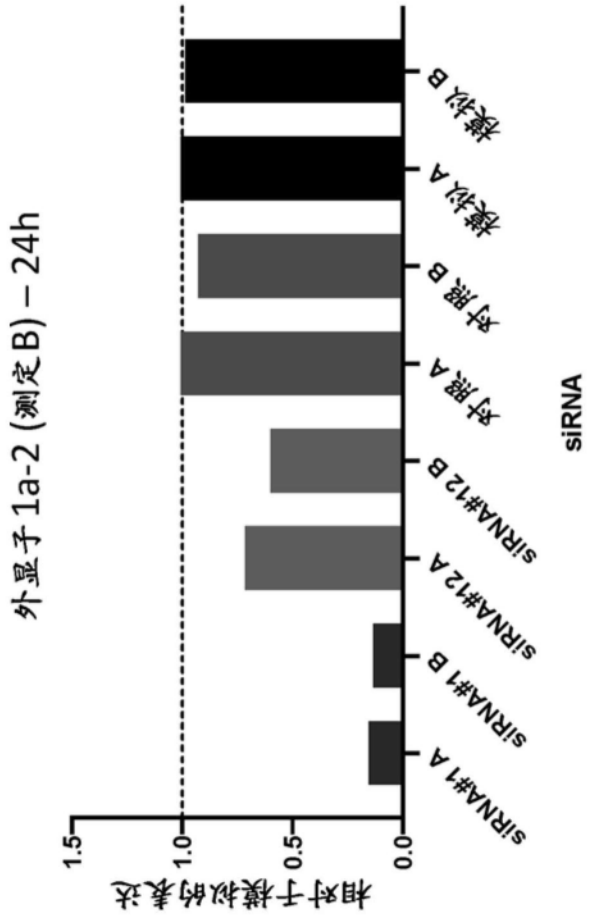

外显子1a-2 (测定B) – 24h
相对于模拟的表达
1.5
1.0
0.5
0.0
siRNA
模拟 B
模拟 A
对照 B
对照 A
siRNA#12 B
siRNA#12 A
siRNA#1 B
siRNA#1 A

图22C

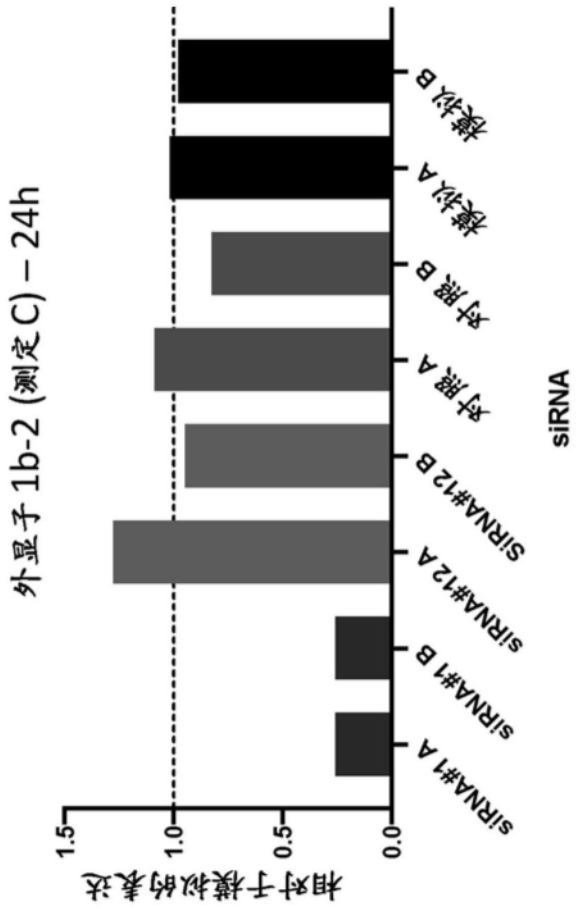

外显子1b-2 (测定C) – 24h
相对于模拟的表达
1.5
1.0
0.5
0.0
siRNA
模拟 B
模拟 A
内照 B
内照 A
SIRNA#12 B
SIRNA#12 A
SIRNA#1 B
SIRNA#1 A

图22D

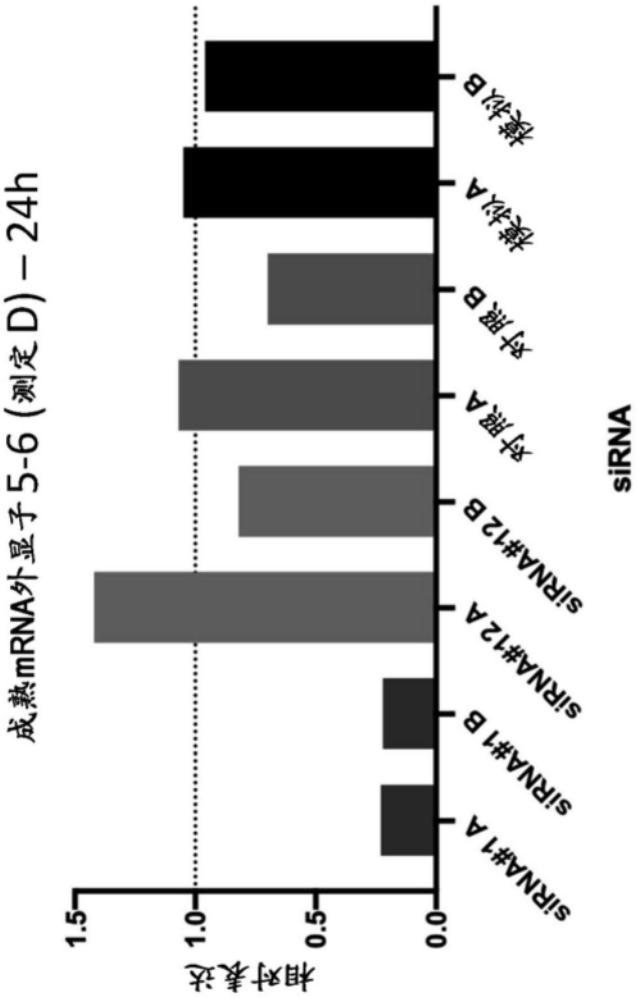

成熟mRNA外显子5-6 (测定D) – 24h
相对表达
1.5
1.0
0.5
0.0
siRNA
阴性对照A
阴性对照B
siRNA#1 A
siRNA#1 B
siRNA#12 A
siRNA#12 B
阳性对照A
阳性对照B

图22E

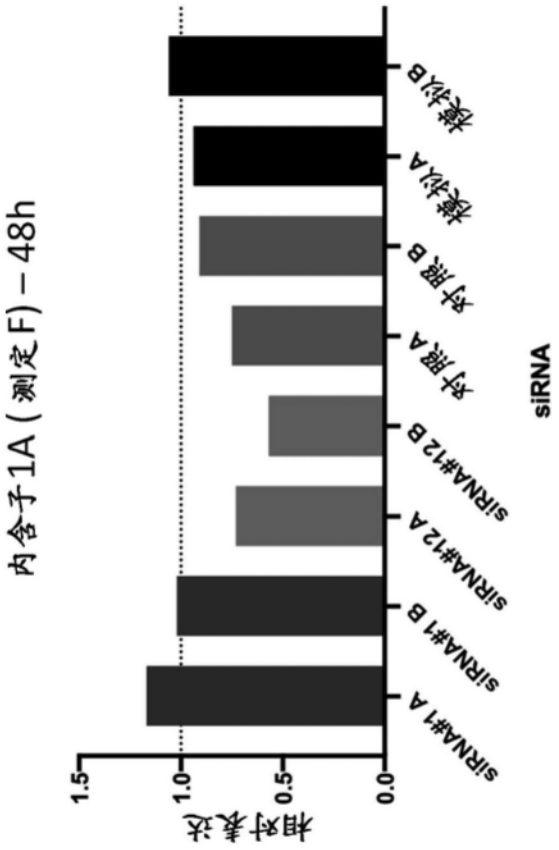

内含子1A（测定F）－48h
相对表达
1.5
1.0
0.5
0.0
siRNA
模拟B
模拟A
对照B
对照A
siRNA#12 B
siRNA#12 A
siRNA#1 B
siRNA#1 A

图23A

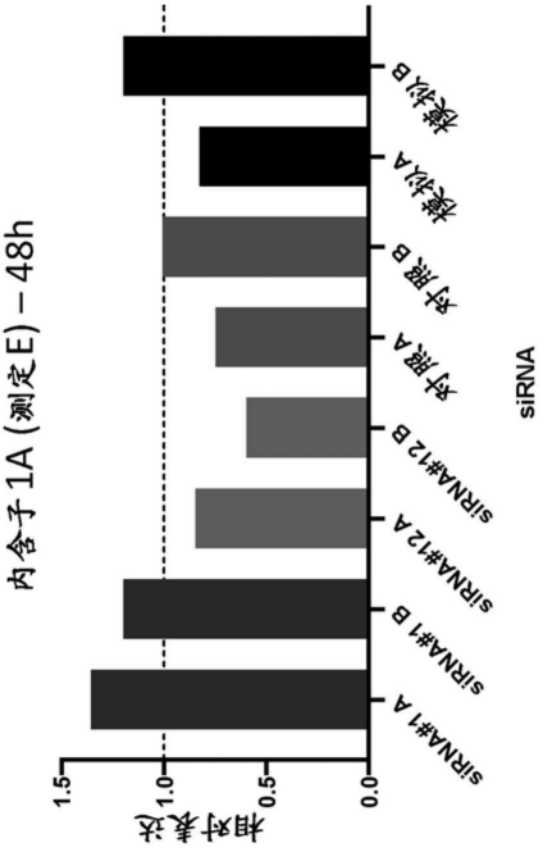

内含子1A (测定E) – 48h
相对表达
1.5
1.0
0.5
0.0
siRNA
siRNA#1 A
siRNA#1 B
siRNA#12 A
siRNA#12 B
对照A
对照B
模拟A
模拟B

图23B

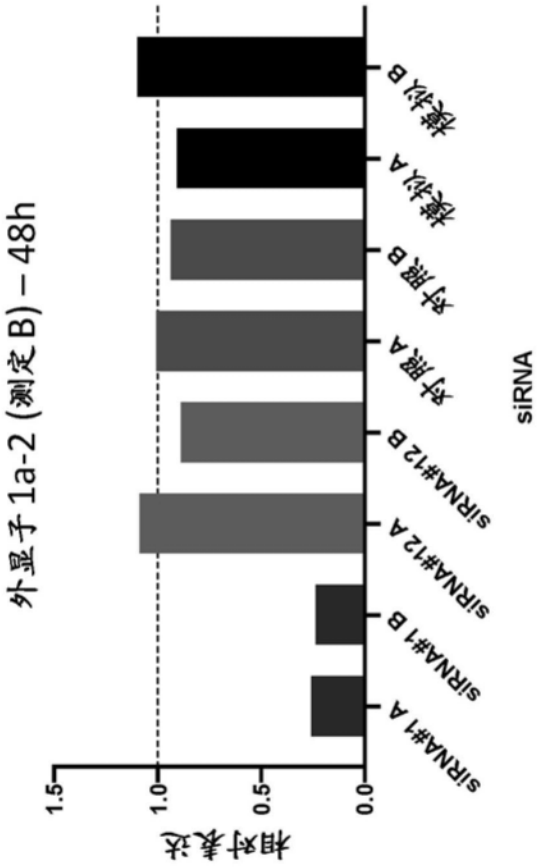

外显子1a-2 (测定B)－48h
相对表达
1.5
1.0
0.5
0.0
siRNA
模拟 B
模拟 A
针对 B
针对 A
siRNA#12 B
siRNA#12 A
siRNA#1 B
siRNA#1 A

图23C

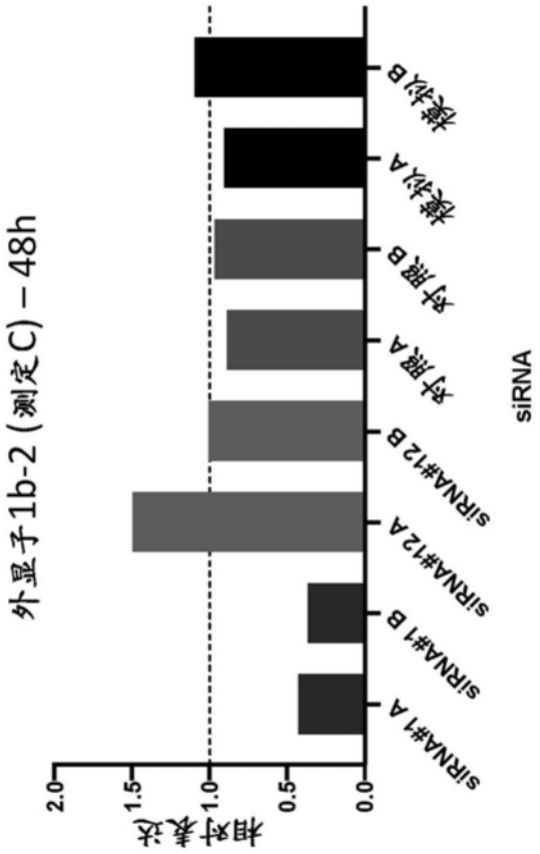

外显子1b-2 (测定C) – 48h
相对表达
2.0
1.5
1.0
0.5
0.0
siRNA
模拟 B
模拟 A
对照 B
对照 A
siRNA#12 B
siRNA#12 A
siRNA#1 B
siRNA#1 A

图23D

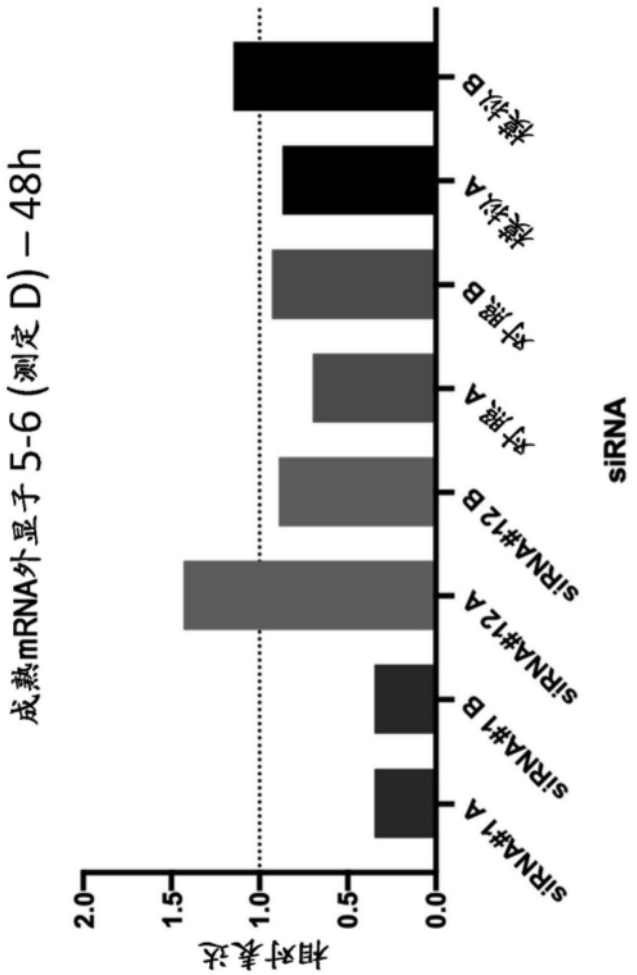

成熟mRNA外显子 5-6 (测定 D) – 48h
相对表达
2.0
1.5
1.0
0.5
0.0
siRNA
模拟 B
模拟 A
对照 B
对照 A
siRNA#12 B
siRNA#12 A
siRNA#1 B
siRNA#1 A

图23E